

国家出版基金资助项目

庞加莱　幼时因智商测验分数之低被列为弱智，成年时又因数学天分之高被誉为最后一个通才。

佩雷尔曼　昔日国际数学奥林匹克金牌选手，今天举世瞩目的俄罗斯数学隐士。

影响数学世界的猜想与问题

佩捷　李莹英　郭梦舒　编著

从庞加莱到佩雷尔曼
——庞加莱猜想的历史

From Poincaré to Perelman
—The History of Poincaré Conjecture



哈尔滨工业大学出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

From Poincaré to Perelman
—The History of Poincaré Conjecture

一个超级猜想　两位旷世奇才
三路人马逐鹿　四方震惊热议

影响数学世界的猜想与问题

- 从费马到怀尔斯——费马大定理的历史
- 从庞加莱到佩雷尔曼——庞加莱猜想的历史
- 从切比雪夫到爱尔特希——素数定理的历史
- 从高斯到盖尔方特——虚二次域的高斯猜想
- 从库默尔到朗兰兹——朗兰兹猜想的历史
- 从比勃巴赫到德布朗斯——比勃巴赫猜想的历史
- 从麦比乌斯到陈省身——麦比乌斯变换与麦比乌斯带
- 从布尔到豪斯道夫——布尔方程与格论漫谈
- 从开普勒到阿诺德——三体问题的历史
- 从华林到华罗庚——华林问题的历史

刘培杰 数学工作室

ISBN 978-7-5603-3916-0



策划编辑　刘培杰
　　　　　张永芹
责任编辑　张永芹
　　　　　王勇钢
封面设计　孙茵艾

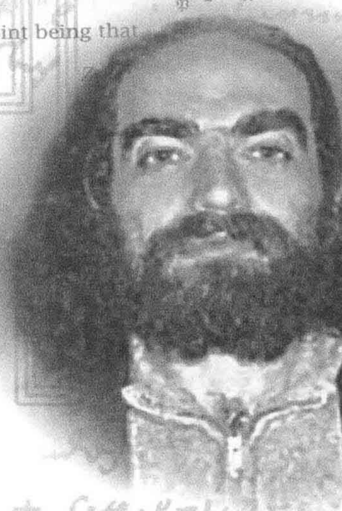
上架建议：科普类／数学史／拓扑学

定价　298.00元



国家出版基金资助项目





影响数学世界的猜想与问题

佩捷 李莹英 郭梦舒 编著

从庞加莱到佩雷尔曼
——庞加莱猜想的历史



rom Poincaré to Perelman
—The History of Poincaré Conjecture



哈爾濱工業大學出版社
HARBIN INSTITUTE OF TECHNOLOGY PRESS

内容简介

全书共分3编23章:上编庞加莱与庞加莱猜想;中编三维空间与拓扑学;下编面向大众的拓扑学描述.详细阐述了庞加莱猜想从提出到解决的全过程以及相关的数学专业理论.

本书适合于高等学校数学及相关专业师生使用,也适用于数学史爱好者.

图书在版编目(CIP)数据

从庞加莱到佩雷尔曼/庞加莱猜想的历史/佩捷,李莹英,郭梦舒编著.—哈尔滨:哈尔滨工业大学出版社,2013.3

(影响数学世界的猜想与问题)

ISBN 978-7-5603-3916-0

Ⅰ.①从…　Ⅱ.①佩…②李…③郭…　Ⅲ.①庞加莱猜测　Ⅳ.①O189

中国版本图书馆 CIP 数据核字(2012)第315002号

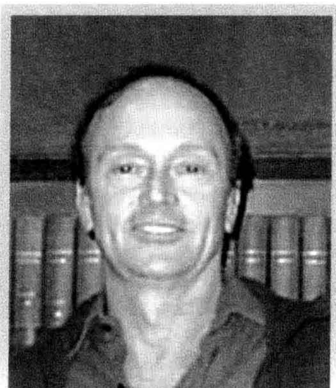
策划编辑　刘培杰　张永芹
责任编辑　张永芹　王勇钢
封面设计　孙茵艾
出版发行　哈尔滨工业大学出版社
社　　址　哈尔滨市南岗区复华四道街10号　邮编150006
传　　真　0451-86414749
网　　址　http://hitpress.hit.edu.cn
印　　刷　黑龙江省教育厅印刷厂
开　　本　787mm×1092mm　1/16　印张72.25　字数1330千字
版　　次　2013年3月第1版　2013年3月第1次印刷
书　　号　ISBN 978-7-5603-3916-0
定　　价　298.00元



亨利·庞加莱（法国数学家）

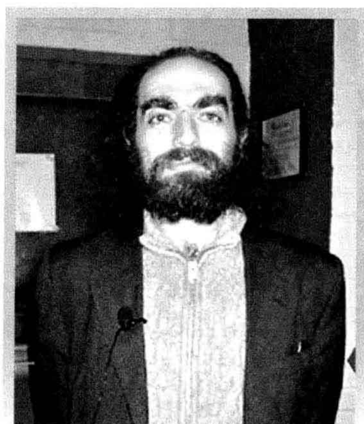






理查德·哈密尔顿

（Ricci流理论的创立者、美国科学院院士）



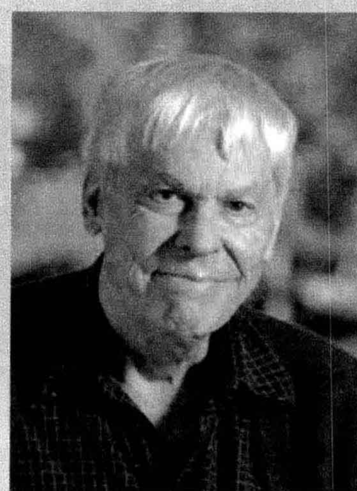
格里戈里·佩雷尔曼

（俄罗斯数学家）



威廉·瑟斯顿

（康奈尔大学教授）



斯蒂芬·斯梅尔

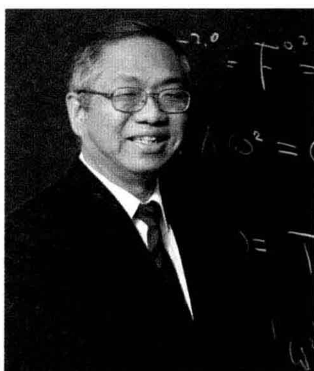
（美国科学院院士）





哈里·福斯坦柏

（美国科学院院士）







丘成桐 （菲尔兹奖获得者）



朱熹平教授（左）曹怀东教授（右）



朱熹平教授（中山大学教授）





曹怀东

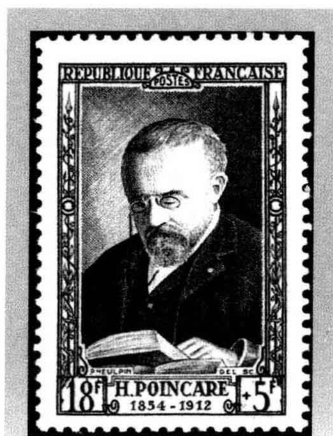
（美国里海大学教授）





田刚

（美国哈佛大学教授）



法国1952年发行的

纪念庞加莱邮票



法国数学家庞加莱

（1854—1912）



波兰数学家巴拿赫

（1892—1945）



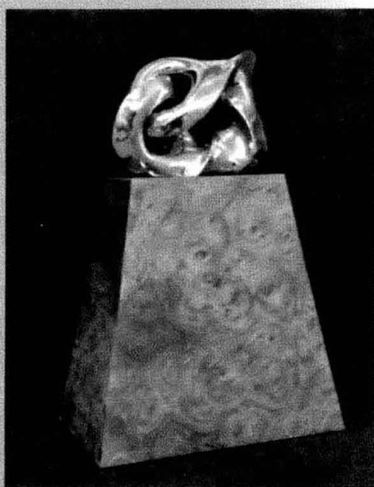
瑞典奥斯卡二世国王



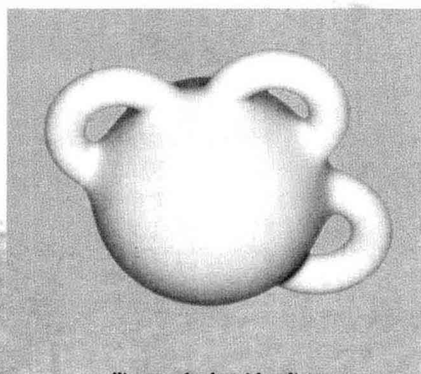
低维拓扑学研究的领袖人物之一瑟斯顿

瑟斯顿在2010年的
巴黎时装秀





拓扑结构的雕塑



带三个柄的球面

菲尔兹奖

菲尔兹奖（Fields Medal，全名 The International Medals for Outstanding Discoveries in Mathematics）是一个在国际数学联盟的国际数学家大会上颁发的奖项。每四年颁奖一次，颁给有卓越贡献的年轻数学家，每次最多四人得奖。得奖者须在该年元旦前未满四十岁。它是据加拿大数学家约翰·查尔斯·菲尔兹的要求设立的。菲尔兹奖被视为数学界的诺贝尔奖。



奖章由加拿大雕塑家罗伯特·泰特·麦肯齐(Robert Tait McKenzie)设计。

正面有古希腊科学家阿基米德右侧头像。在头像旁刻上希腊文“Α Ρ Χ Ι Μ Η Δ Ο Υ Σ”，意思为“阿基米德的（头像)”。又刻上作者名字缩写 RTM，和设计年份的罗马数字 MCNXXXIII（1933 年，第二个 M 字以 N 代替），还有一句拉丁文“TRANSIRE SUUM PECTUS MUNDOQUE POTIRI”，意为“超越他的心灵，掌握世界”，出自罗马诗人马尔库斯·马尼利乌斯(Marcus Manilius)的著作《天文学》(Astronomica)卷四第 392 行。句中“suum”（他的）原文作“tuum”（你的）。

奖章背面刻有拉丁文“CONGREGATI EX TOTO ORBE MATHEMATICI OB SCRIPTA INSIGNIA TRIBUERE”，意为“聚集自全球的数学家，为了杰出著作颁发（奖项）”。背景为阿基米德的球体嵌进圆柱体内。

得奖列表

年份	地点	得主（国籍，当年生日后岁数）
1936	挪威奥斯陆	拉尔斯·瓦莱里安·阿尔弗斯（芬兰，29岁）
		杰西·道格拉斯（美国，39岁）
1950	美国麻省剑桥	洛朗·施瓦茨（法国，35岁）
		阿特勒·塞尔贝格（挪威，33岁）
1954	荷兰阿姆斯特丹	小平邦彦（日本，39岁）
		让-皮埃尔·塞尔（法国，28岁）
1958	英国苏格兰爱丁堡	克劳斯·弗里德里希·罗斯（英国，33岁）
		勒内·托姆（法国，35岁）
1962	瑞典斯德哥尔摩	拉尔斯·赫尔曼德（瑞典，31岁）
		约翰·米尔诺（美国，31岁）
1966	俄罗斯莫斯科	迈克尔·阿蒂亚（英国，37岁）
		保罗·寇恩（美国，32岁）
		亚历山大·格罗滕迪克（无国籍，居于法国，生于德国，38岁）
		斯蒂芬·斯梅尔（美国，36岁）
1972	法国尼斯	艾伦·贝克（英国，31岁）
		广中平祐（日本，39岁）
		谢尔盖·诺维柯夫（苏联，32岁）
		约翰·格里格斯·汤普森（美国，38岁）
1974	芬兰赫尔辛基	恩里科·邦别里（意大利，34岁）
		大卫·芒福德（美国，生于英国，37岁）
1978	加拿大温哥华	皮埃尔·德利涅（比利时，34岁）
		查尔斯·费弗曼（美国，29岁）
		格列戈里·亚历山德罗维奇·马尔古利斯（苏联，32岁）
		丹尼尔·格雷·奎林（美国，38岁）
1982	波兰华沙	阿兰·孔涅（法国，35岁）
		威廉·瑟斯顿（美国，36岁）
		丘成桐（美国，生于中国，33岁）

得奖列表

年份	地点	得主（国籍，当年生日后岁数）
1986	美国加州伯克利	西蒙·唐纳森（英国，29岁）
		格尔德·法尔廷斯（西德，32岁）
		迈克尔·哈特利·弗里德曼（美国，35岁）
1990	日本京都	弗拉基米尔·德林费尔德（苏联，36岁）
		沃恩·弗雷德里克·兰德尔·琼斯（新西兰，38岁）
		森重文（日本，39岁）
		爱德华·威滕（美国，39岁）
1994	瑞士苏黎世	让·布尔甘（比利时，40岁）
		皮埃尔–路易·利翁（法国，38岁）
		让–克里斯托夫·约科兹（法国，37岁）
		叶菲姆·泽尔曼诺夫（俄罗斯，39岁）
1998	德国柏林	理查德·博赫兹（英国，生于南非，39岁）
		威廉·蒂莫西·高尔斯（英国，35岁）
		马克西姆·孔采维奇（俄罗斯，34岁）
		柯蒂斯·麦克马伦（美国，40岁）
2002	中国北京	洛朗·拉佛阁（法国，36岁）
		弗拉基米尔·沃埃沃德斯基（俄罗斯，36岁）
2006	西班牙马德里	安德烈·欧克恩科夫（苏联 美籍，37岁）
		格里戈里·佩雷尔曼（俄罗斯，40岁）
		陶哲轩（澳大利亚，31岁）
		温德林·沃纳（法国，38岁）
2010	印度班加罗尔	吴宝珠（法国 越南裔，38岁）
		埃隆·林登施特劳斯（以色列，40岁）
		斯坦尼斯拉夫·斯米尔诺夫（俄罗斯，40岁）
		塞德里克·维拉尼（法国，37岁）

⊙前言

沙漠植树是一个善举.

数学史学家胡作玄先生在《近代数学史》(济南:山东教育出版社,2006年)的序言中指出:

世界数学史的研究有资料的限制,语言的困难和文化背景的差异.特别地,近代数学史的研究更要求研究者具备数学的功底,科学的知识及哲学的头脑,没有一定的训练恐怕难有什么成就.如果再加上社会上的反智传统,科学界对于历史的无知及蔑视,那么其后果就只能是呈现出一种文化沙漠的景象.

在此文化沙漠上植树种草,不论活与不活都是值得一试的,更何况是一株法国梧桐.

英国《金融时报》刊登的专栏作家西蒙·库珀的文章说:"人们过去常说每个文明人都有两个祖国:一个是自己的国家,另一个是法国."

对于数学史来说这句话太准确了.法兰西这个以浪漫闻名于世的民族居然在以理性与严谨为其特征的数学中屡屡领先.庞加莱这位曾被著名的IQ测试判为"傻子"的法国人近年来再一次成为世界关注的焦点.原因是他提出的著名庞加莱猜想被俄国数学家佩雷尔曼证明了.

美国数学家阿诺德·塞庚在谈及法国的数学时,用一首叫"至美"的诗对法国数学,特别是庞加莱进行了赞美,这首诗是这样的(翻译成中文后已无任何美感)

天皇皇,地皇皇.
庞加莱,真叫棒.
证明漂亮定理,
不慌也不忙.
使用对偶,还有亏格,
就把流形微分同胚分类光.

德国著名导演赫尔佐格导演了一部颇为另类的关于南极的电影,片名叫《在世界的尽头相遇》,在这部片子中最让人难忘的是那只决绝的企鹅,它从队伍中转头,背离海洋向70公里外的群山狂奔,“它的步态滑稽而孤独,癫狂而决绝,那确实是难以忘怀的背影.”

有人说:每个物种都有这样的“开拓者和好奇者”.佩雷尔曼恰似人类中那只决绝的企鹅.

法国的纪德说:“别人比成功,我愿比永久.”

本书所记录下的几百位在征服庞加莱猜想征程上的世界各国数学家因他们各自的出色工作,虽然没有取得最后的成功,但都永久地镌刻在人类智库的金榜上.这其中也有我们中国数学家的名字,尽管在国际舆论界的承认问题上有许多戏剧性的情节,但这毕竟是在世界级的著名数学猜想的解决过程中闪现出了国人的身影,这是足以令我们感到自豪的.

2005年11月,中科院数学与系统科学研究院举办“华罗庚讲座”学术报告会,邀请了对庞加莱猜想作出开创性贡献的美国科学院院士、数学家哈密尔顿作学术报告,他在黑板上写下了三四位华人的名字,其中就有朱熹平和曹怀东,还提到了他们的工作.

数学界与围棋界颇类似,都是靠实力、靠硬功夫打天下,来不得半点虚假,但正如中、日、韩三足鼎立.要想分出绝对高下,也并非易事.

有人这样写道:万涓成水,岁月铅华洗净,拨开数十载乱云尘土,整个围棋世界才发现唯有吴清源和李昌镐笑对夕阳黄昏,无愧无咎,参天而立……

在证明庞加莱猜想的数十载风云变幻中最后参天而立的两棵大树是哈密尔顿和佩雷尔曼,这一点在国际数学界似已有定论,而且这两位都是颇具传奇色彩的人物,与国人想象中的数学家及现实生活中的数学家都很不相像.

哈密尔顿(Richurd Hamilton)是国际公认的几何分析大师.据说在证明庞加莱猜想的功劳簿上他应该分享50%以上,但他不是传统意义上刻板的数学家,而是颇有浪漫气息的人,据曹怀东的夫人映碧在《偶然尘》中描述:“有一次,他一个人在家生火,炉火不知怎么就烧到炉壁外面来了,急得这位天才数学家赶紧给房东老太太打电话,借灭火器.没想到这一回电话那边传来的却是房东侄

女的甜美声音.噢,你们知道,她可是个金发靓女!在这之前Richurd正愁没得着机会跟她唠嗑呢,屋里浓烟弥漫,Richurd仍旧与靓女聊了近一刻钟,然后才问她:'我能借你的灭火器用一下吗?''干嘛?''我屋里着火了.'那边自然是啊的一声惊叫."(映碧,著.偶然尘.长沙:湖南科学技术出版社,2010:273)

就是这样一位颇具浪漫情怀的数学家才会想到用瑞奇流来攻克庞加莱猜想,这一大胆的想法连以敢想著称的丘成桐大师都感到"是疯了".哈密尔顿曾在几年前来过一次中国,不过不是来做学术报告的,据说是来追求某位心仪的中国女士的.

北京师范大学教育学部博士生高政在一篇题为"自我为什么是深渊?"的文章中指出:"……在今天看来,有些哲学家自以为成功的思想成果未必是真的成功.但至少,他们自己会感觉幸福而圆满.有一件自己喜欢又毕生追求的事情是多么的美好,如果碰巧解决了,就像牛顿谦虚地说他在海边捡到几个贝壳一样,真是要感谢上帝了!今天,有一件自己喜欢又毕生追求的事情,几乎成为大多数人的奢望."(张立升,主编.社会学家茶座.济南:山东人民出版社,2011年第2辑)

本书虽然是讲述拓扑学中的一个重要猜想,但多数读者可能对隐士佩雷尔曼和骑士哈密尔顿及斗士丘成桐更感兴趣.他们都有自己独特的人生观、价值观、世界观.正如哈佛大学的阿瑟·杰夫所说:"我觉得他(佩雷尔曼)是一个非常反传统的人,他拒绝与炫耀和偶像崇拜扯上关系."

佩雷尔曼的与众不同归根到底是他是一个有信仰的人.中国社会科学院副研究员海青对此有一番高论:

信仰容易让人吃亏犯傻,甚至丧命,没有信仰的人应该活得更舒服.但是,信仰是货真价实的双刃剑,有它没它都很痛苦.如果一个人什么都没信仰过,没为钱以外的任何东西,自己以外的任何人牺牲过哪怕一根头发,甚至买东西讨价还价都没吃过一次亏,那么他到中年时,可能会感到空虚莫名,上蹿下跳想找到一点儿牺牲的感觉,这种"信仰型"中年危机,比什么都无药可救.

这本书从广义上说可以算做拓扑学史书.作家石康曾说:"我想要干点实事儿,不怕别人笑话,我把它说出来,那就是,我要写一本浅显易懂又生动有趣的《数学史》.这样的书在国内没有(早就有,但他不知道),我要写出它,也可让后面有天赋的中国小孩对数学发生兴趣,要是万一有人能干出比高斯、牛顿、阿基米得更有价值的事来,那么我也可以得到少许安慰.当然,我就是再不切实际,也不会荒唐到指望中国能出个伯努利家族.我知道,那不可能,但我仍存侥幸心理,希望个别天才能横空出世,对世界有真正持久的贡献."

现在天才终于出世了,但不是中国的朱和曹,而是佩雷尔曼.奇怪的是

2003年当俄罗斯斯捷克洛夫数学所没有推选佩雷尔曼为其研究员时，他的信心遭到了打击，据他的一位朋友说，他感到自己是一个“无能而缺乏天赋的人”．尽管他早在16岁那年即在1982年国际数学奥林匹克竞赛中以满分获得金奖，而且他还是一位颇具才华的小提琴手和乒乓球选手．

本书与一般的科普书相比难度超大，但非刻意而为，牛顿之所以要把《自然哲学的数学原理》写得很晦涩难懂，以至有人将其称为“最难看懂的书之一”，是因为他害怕被那些他所认定的数学“门外汉”纠缠不休．

现代的中国也是如此，“民科”遍地，只要有某一个世界难题他们一知半解，便开始大肆向媒体与出版机构进行所谓已证明的狂轰滥炸，所以建立一定的阅读门槛虽然会失去一部分虽对数学感到好奇但缺少必要数学修养的读者，但得到的好处是进行了有效与精准的传播．

本书其实是对庞加莱猜想获证这一重大历史事件的一个全方位解读，包括历史、文化、哲学、心理学等诸多方面，在此我们声明，我们能做的仅是解读而已．

我们既像狗仔队到处搜寻有关这一猜想的“涉案”学术明星的点滴个人信息，也像影评人一样按照自己的一知半解将煽动起观众热情．这种工作是需要的，否则像庞加莱那样生活在上个世纪初的法国贵族和像佩雷尔曼那样生活在现代的俄罗斯隐士是很难被了解的．佩雷尔曼低调到了极致，正如纽约州立大学数学家迈克尔·安森所说：“佩雷尔曼来过了，解决了问题．其他的一切对于他都是肤浅的．”如果没有人像我们这样关注数学家那炷幽明的香火和数学儒林那份执著的传承，这伟大的时刻将会被淹没．

在人们对数学史多少有些了解后都会产生这样的疑问：为什么数学家如此倾向于将毕达哥拉斯共济会的高度秘密在他们之中严加保守．对普通人来说，他们“真实的”世界的完善意味着虚幻．普通人生活于其中的世界，乃是一个斗争和失败、实验和错误的世界．而在数学世界中，一旦你习惯于任一事物后，它都是显而易见的，我们通常被蒙在鼓里的是：人类也许费了一千年时间，才看出了在数学论证中只需举手之劳就“显而易见的”事物．

1976年一位英国文人威尔逊送给董桥一本《The Colossus of Maroussi》时说了一句：“世界太喧闹了，我们差一点错过了这样远古的一声喟叹！”对此我们深有同感．

刘培杰

2012年5月27日

目录

上编 庞加莱与庞加莱猜想

中编 三维空间与拓扑学

上　编

庞加莱与庞加莱猜想



庞加莱猜想获证

引

言

1 庞加莱猜想的历史与解决[①]

20世纪末,数学家怀尔斯(Andrew Wiles)解决数论上高悬三百多年的重要猜想——费马最后定理.最后,一个拓扑学上最重要的问题——庞加莱猜想(Poincaré Conjecture),也在提出近百年后,在21世纪初被攻克了.俄国数学家佩雷尔曼(Grisha Perelman)和美国数学家汉米尔顿(Richard Hamilton)是解决庞加莱猜想的最大功臣.

2006年8月底,四年一度的国际数学家大会(ICM)在西班牙首都马德里隆重举行,庞加莱猜想的验证,是这场盛会中最令人振奋及众人的焦点所在.正如同大家所预期的,佩雷尔曼获得了素有数学界诺贝尔奖之称的菲尔兹奖(Fields Medal),这个奖每四年才颁发一次,而且得奖者年龄必须在40岁以下.然而,佩雷尔曼不仅没有出席本次的国际数学家大会,还史无前例地拒绝接受菲尔兹奖,引起众人许多错愕及遗憾.

庞加莱猜想,是法国数学家庞加莱(Henri Poincaré)在1904年所提出的一个关于三维流形的重要猜想.这个猜想是20世纪带动拓扑学研究及发展的主要动力,百年来数学家前仆后继,不屈不挠,在不断尝试的过程中,逐渐累积了我们对拓扑学更深一层的了解.

世纪之交,庞加莱猜想依旧名列数学上最重要的未解问题之一,美国数学家,同时也是菲尔兹奖得主的斯梅尔(Stephen Smale),在1997年将它纳入“21世纪18个重要数学问题”,这个猜想也是美国克莱数学研究所(Clay Mathematics Institute)悬赏百万美金的七大名题之一.

---

① 原作者李莹英.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

庞加莱猜想的解决,依赖三十多年来在几何分析领域中所发展及累积的知识及技巧,而为解决这个问题所发展出的处理奇异点的新方法,也极有可能会在新世纪中,对研究其他具奇异点的重要问题,产生突破性的影响。

数学家们在10年中目睹了两个史上著名难题接续被解决,不禁要臆测是否数学界将进入一个百花齐放的黄金时期,我们正热切期盼着这个新纪元的到来.在这个令人振奋的历史时刻,本文希望将庞加莱猜想的背景及解决过程,作一个粗浅的介绍,以飨所有的读者.

一、什么是庞加莱猜想?

要介绍庞加莱猜想,首先必须认识拓扑学(Topology).数学家研究各式各样的数学结构,并试图在这些结构中作完整的分类.所谓的拓扑学,就是将能像橡皮糖一样变形得到的东西(不能扯断),都视为一样.所有的二维曲面(设为可定向、有限且没有边界),都可以变形成球、一个洞、二个洞、三个洞等(图1);也就是说,二维曲面可以用它有几个洞来完全刻画.接下来最重要的问题,就是对三维空间作分类,所谓的三维是指有三个自由变动的独立方向.



图1

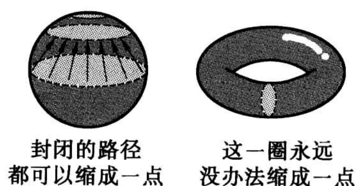


图2

我们可以从所处的三度空间中“看到”一个二维曲面的样子,但对于生长在曲面上的蚂蚁,它却无法飞出曲面来看看自己所生存的曲面.它是否能在不离开曲面的情况下,就知道自己的曲面是哪一种拓扑形态呢?答案是可以的.例如二维球具有一个性质,任何一圈路径都可以在曲面上缩成一个点,这在其他有洞的曲面是办不到的(图2).

人类在面对所生活的宇宙时,也和蚂蚁处于相同的困境,我们无法脱离所在的“这个世界”,从外面去“看到”它长什么样子.那我们是否能像二维曲面上的蚂蚁一样,从所在的“这个世界”中,推测我们所处的空间的模样呢?

这项工作,有赖我们对三维空间的了解及分类,其中最基本的问题就是庞加莱猜想.庞加莱仿照二维球定义三维球$\{(X, Y, Z, W) \mid X^2+Y^2+Z^2+W^2=1\}$,这可以说是最简单的三维有限且无边界的空间.本文所探讨的空间,都将局



限在有限且无边界的空间,数学上称这个空间为“紧致空间”.

庞加莱所提出的问题是:“一个紧致的三维空间,如果它上面每一条封闭曲线都可以缩成一个点,则这个空间是否必定和三维球相同?”叙述里的相同,指的就是我们在前面所介绍的“拓扑”相同,亦即可以透过橡皮糖变形成完成一样.这个问题后来就被大家称为庞加莱猜想.

二、解决猜想的漫漫长路

数学家在一开始面对这个问题时,忽略了其中一些细腻的复杂现象及困难,曾经有许多知名的数学家提出过错误的证明.事实上,第一个错误即是庞加莱自己所犯.他在1900年以为自己证得一个更广的定理,后来他自己发现了反例,才修正成目前的叙述.虽然走了许多冤枉路,但是每次的错误都让我们更进步一些.数学家在发现证明的错误及寻找反例当中,逐步建立及累积我们对三维拓扑的知识.

“一般化”是数学家研究问题时常用的一种策略,因此数学家试着探讨一般维度的庞加莱猜想,希望能一举解决这个问题.一般人可能以为如果维度越高,问题就会更困难,但事实上并不一定如此.1956年美国数学家米尔诺(John Milnor)发现,如果将上述的“相同”解释成保持微分结构的相同,那七维以上将会有不只一个球.这表示研究庞加莱猜想时,必须明辨“相同”的意义.米尔诺以这个震惊当时数学社群的结果,获得了1962年的菲尔兹奖.

没隔多久(在1960年),另一位美国数学家斯梅尔就以一贯的方法,证明了五维以上的庞加莱猜想,他也因这项贡献而获颁1966年的菲尔兹奖.但很可惜的是斯梅尔的一般方法在低维度的三维和四维并不适用,因此并没有解决原来的问题.

四维的情况,要等到约20年后才有突破性的新发展.1982年美国数学家福里德曼(Michael Freedman)利用两个不变量,对某一大类较简单的四维紧致拓扑空间作了分类,他的分类结果特别地解决了四维的庞加莱猜想.而在福里德曼稍后,英国数学家多纳森(Simon Donaldson)利用偏微分方程,做了一些四维可微流形拓扑限制的深刻工作,结合两人的工作,甚至证明了普通四维欧氏空间上有无穷多个可微结构的结果.这些令人兴奋的结果,让多纳森及福里德曼获颁1986年的菲尔兹奖.

当然,也有很多数学家谨守研究本来三维的庞加莱猜想,在三维拓扑的研究不断地得到许多深入的结果.其中最重要的一个成就,来自另一名美国数学家瑟斯顿(William Thurston).

瑟斯顿在1978年针对三维拓扑的分类,刻画出一个细致而完整的研究方略,这个涵盖了庞加莱猜想的想法,后来被称为几何化猜想(Geometrization Con-

jecture).他认为所有的三维拓扑空间,都是由一些基本类型的空间再以特殊的方法结合组成,而这些类型的空间之所以称为基本,是因为上面都有很好的距离度量,这也正是“几何化”得名的由来.

几何化猜想带动了另一波新的研究风潮.然而瑟斯顿自己与其他数学家所证得的许多结果,都是关于负曲率的情况(这相当于二维曲面中洞的个数大于1的情况),对原始的庞加莱猜想依然束手无策.瑟斯顿因为他对三维空间的开创性想法,而获颁1982年的菲尔兹奖.

三、看见曙光——瑞奇流

在瑟斯顿的工作之后,数学家不再对庞加莱猜想是否正确有所怀疑,开始有数学家假设庞加莱猜想正确,然后继续做进一层的探讨及研究.而另一方面,由于瑟斯顿预测在三维的基本类型上,均存在好的距离函数,这也意味着可能可以用分析的办法来解决这个拓扑的问题.谈到几何中的分析方法,就不能不提另外一个数学家——丘成桐.

丘成桐是1982年的菲尔兹奖得主,也是第一位获颁这个奖项的华裔数学家.从1970年代中期起,当时才20几岁的丘成桐与一群合作者,将非线性分析的方法及技巧大量地引进几何问题的研究,并接连解决许多著名的猜想及重大结果,这也开创了一个现今非常重要的数学领域——几何分析(Geometric Analysis).庞加莱猜想的解决,所依赖的正是这三十多年来所发展的几何分析知识.

1979年,丘成桐和孙理察(Richard Schoen)首先利用极小曲面的分析理论来研究三维拓扑.佩雷尔曼及汉米尔顿分析瑞奇流(Ricci Flow)的退化行为这个重要问题时,丘成桐和孙理察的上述结果居于关键的地位.

丘成桐对庞加莱猜想的相关研究,一直非常注意.在瑟斯顿的几何化猜想之前,研究庞加莱猜想的主要拓扑工具,是迪恩引理(Dehn's Lemma)、回路定理(Loop Theorem)以及球定理(Sphere Theorem).1981年,密可斯(William Meeks)和丘成桐利用极小曲面的分析方法,不但重新证明了这三个结果,同时更推广到有限群作用下的等价观点,他们也因此解决一个知名的拓扑问题——史密斯猜想(Smith Conjecture).

这些都为利用几何分析来研究拓扑问题的威力提供有利的佐证,更进一步与庞加莱猜想密切相关的,是汉米尔顿在1982年所提出的瑞奇流.所谓瑞奇流,就是让距离函数随着时间往2倍瑞奇曲率(Ricci Curvature)的相反方向变动.这很像热传导的方式,虽然一开始的温度并不一定均匀,但随着时间的累积,温度的变化会遵循热传导方程式,而逐渐地均匀化,使得各处的温度趋于相同.

我们期望瑞奇流同样能将距离函数逐渐均匀化,最后变成处处相同的常曲





率距离度量,这是非常自然的想法,但其中的困难很多.瑞奇流与热传导方式不同,不只牵涉一个函数,而是许多函数,或者更正确地说是牵涉所谓的张量,这是一个复杂的非线性方程组;随着时间的变动,可能很快就会产生奇异点,接下来要怎样继续变动、怎样控制,全部是困难.在汉米尔顿 1982 年的文章中,他证明当瑞奇曲率是正的时候,整件事是行得通的.丘成桐对汉米尔顿的瑞奇流研究,非常关注及感兴趣,他向汉米尔顿建议可以用瑞奇流来做庞加莱猜想.

丘成桐与李伟光在 1986 年发表了关于热传导方程式一些哈纳克不等式(Harnack Inequality)的先验估计;丘成桐建议汉米尔顿在瑞奇流中研究及推广相应的结果,他相信这绝对会产生很重要的影响.

几年后,汉米尔顿果然将这个结果推广到张量场中,这是一个非常重要的结果,但距离解开庞加莱猜想还相当远,因为还得要分析流形是怎么变化.1995 年汉米尔顿发表一篇文章,讨论在瑞奇流中如何控制奇异点的形成,如何做“手术”(Surgery)及流形的分解,使瑞奇流能继续进行.这往证明庞加莱猜想又迈前一大步,解决这个世纪谜团似乎已经指日可待了,但眼前依然有几个重大困难尚待克服.首先,我们是否能控制所需手术的次数? 再者,经手术动作的瑞奇流的解,其几何及长时间的行为是否都能被充分了解及控制?

四、突破瓶颈

从 1982 年汉米尔顿提出瑞奇流后,这个领域就开始蓬勃发展,丘成桐有许多学生,如曹怀东、周培能、施皖雄等都对这个领域做过非常重要的贡献.近几年,中国中山大学的朱熹平及陈兵龙,也获得了许多非常杰出的结果.瑞奇流的研究需要用到许多估计及分析,是个不容易入门且具高度技术的领域.

2002 年 11 月佩雷尔曼在网络上张贴一篇文章,提出一些开创性的想法来解决汉米尔顿所遭遇到的困难.他的作法若被证实,将可以顺利完成汉米尔顿所起始的,利用瑞奇流来解决庞加莱猜想及几何化猜想的计划.

这个消息不但震撼了从事瑞奇流研究的数学家,整个数学界甚至一般媒体均十分关注.没有人知道佩雷尔曼在做关于瑞奇流的研究,但他的文章显示出他对瑞奇流相关的文献及结果相当熟悉.虽然大家都对这个消息非常振奋,然而包括瑞奇流的专家,都很难在短时间内完全理解及验证佩雷尔曼的文章内容.

这有几个原因,首先是佩雷尔曼的文章写得相当简洁.他在文章中对证明提供大略证明,以及臆测的想法,有很清楚的区分.然而,即使他认为已经证明的部分,也仅是将重要的步骤及结果提纲挈领地写出来.佩雷尔曼思索这个问题将近十年,有些想法或许对他而言再清楚不过了,但在这段时间内,佩雷尔曼从未跟其他人沟通交流,外界对于这些想法及概念自然不熟悉.

另外一个原因,是有些佩雷尔曼的想法从其他领域引入,这对做瑞奇流的

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

专家而言,是不熟悉的主题,而对这些主题有涉猎的人,则可能完全不了解瑞奇流的相关研究.

佩雷尔曼在前苏联的圣彼得堡大学取得博士学位,1992 年秋天及 1993 年春天分别访问美国纽约大学及石溪大学,1993 ~ 1995 年任教于伯克利大学.1995 年他虽然获得一些美国学校的教职,但选择回到故乡圣彼得堡.

1992 ~ 1993 年普林斯顿高等研究院数学所的主题是几何,包括汉米尔顿的许多知名几何学家都在那里访问.当时笔者也刚毕业,正在高等研究院做博士后研究,我记得每星期佩雷尔曼都会从纽约来听汉米尔顿的课及参加其他活动.他非常害羞,几乎都是独来独往,但当时他已经是个令人瞩目的新秀,很能洞悉问题的本质,并提出简洁的作法.

佩雷尔曼的后续研究,主要是探讨一种流形的极限空间上的性质.正因为这样的背景,所以当 1993 年汉米尔顿在伯克利向他提及利用瑞奇流来作庞加莱猜想和几何化猜想的计划,以及其中所遭遇的一些困难时,他很快就意识到自己在这种极限空间上一个关于塌陷(Collapsing)的结果,可能可以派上用场.

汉米尔顿是个十分具有原创性的数学家,他习惯用自己的想法解决数学问题,所以当佩雷尔曼向他提出自己的观察时,他可能完全没有放在心中.汉米尔顿同时是个热情而且心胸开阔的人,他毫无保留地告诉佩雷尔曼他的想法,包括几年后他才发表的工作(笔者臆测是 1995 年所发表关于奇异点的分析及流形分解).

佩雷尔曼从这些会谈及汉米尔顿文章中学习瑞奇流的知识,所以虽然佩雷尔曼并不直接受教于汉米尔顿,但他自认为是汉米尔顿的门徒.

五、褪尽迷雾,展望未来

佩雷尔曼 2003 年 3 月在网络上张贴另一篇 22 页的文章,以证明上一篇中被省略以及较技术性的部分,但这依然不能理清众人的疑惑.2003 年 4 月他应邀到美国各知名大学演讲访问一个月,演讲中佩雷尔曼解答大家所提的一些问题.佩雷尔曼的这趟旅程显然最想和汉米尔顿见面讨论,但或许整件事的发展太出乎汉米尔顿意料,他们之间并没有什么互动.同年 7 月,佩雷尔曼在网络张贴最后一篇 7 页的文章后,拒绝提供更详尽的版本,同时也不回大部分人的电子邮件.

有些数学家汲取佩雷尔曼文章中的部分想法,应用到其他问题,并且得到很好的进展.然而庞加莱猜想到底解决了没? 谁来背书及验证呢? 这是件困难而且吃力不讨好的事,个人原先的研究必须先放在一边,投注几年的时间,可能最后没有任何收获.而由于无法理解佩雷尔曼的一些想法及所使用的工具,从事瑞奇流研究的数学家,还得尝试用其他的方法来证明及补足这些不清楚的地方.





由于庞加莱猜想是克莱数学研究所悬赏百万美金的七大问题之一,他们于2003年主动委托密西根大学的克莱那(Bruce Kleiner)和洛特(John Lott),以及另一组人马——哥伦比亚大学的摩根(John Morgan)及普林斯顿大学的田刚——来检查及验证佩雷尔曼的工作.

克莱那及洛特陆续在网络上发表检查的进度,2006年5月他们在网络上张贴较完整的手稿;2006年6月,中国中山大学朱熹平及美国里海大学的曹怀东则在《亚洲数学期刊》(The Asian Journal of Mathematics)发表他们的完整证明;2006年7月摩根及田刚也在网络上发表他们的结果.大家一致确认佩雷尔曼的想法是行得通的,因此庞加莱猜想及几何化猜想已经顺利被解决了!

庞加莱猜想的解决,除了结束一个高悬已久的重要问题及世纪难题,更重要的是,我们希望在这个过程中所发展出来的新技巧,能协助我们解决其他重要问题.瑞奇流是一个复杂的非线性偏微分方程组,这是数学家们第一次能够完全理解复杂的偏微分方程组的奇异点和演化.类似的方程组在自然现象中比比皆是,例如爱因斯坦方程以及 Navier Stokes 方程等,解决瑞奇流的方法,将对这些自然的方程的研究带来曙光.另外,其他如四维中可微版本的庞加莱猜想、物理中重整化群的概念,以及均曲率流等,亦是这些新想法可能会派上用场的地方.

1904~2006年庞加莱猜想的解题接力赛

年	数学家对庞加莱猜想的贡献
1904	法国庞加莱提出猜想.
1956	美国米尔勒提出一个相关问题的反例,获颁1962年菲尔兹奖.
1960	美国斯梅尔证明五维以上维度的庞加莱猜想,获得1966年菲尔兹奖.
1978	美国瑟斯顿提出三维拓扑分类一些更细致完整的影像以及几何化猜想,获得1982年菲尔兹奖.
1982	美国福里德曼解决了四维的庞加莱猜想,获得1986年菲尔兹奖;同时,汉米尔顿提出瑞奇流.
1995	汉米尔顿发表文章,讨论瑞奇流如何控制奇异点的形成,以及手术与流形的分解,使瑞奇流顺利进行.
2002	俄国佩雷尔曼在网络张贴第一篇文章,提供关于几何化猜想的概略证明,随后在2003年3月及7月贴出另两篇.
2006	美国克莱那和洛特张贴验证佩雷尔曼证明的较完整手稿.
2006	中国朱熹平和美国曹怀东在亚洲数学期刊发表论文,表示完整证明了庞加莱猜想及几何化猜想.
2006	美国摩根和田刚也张贴他们验证佩雷尔曼证明的文章.
2006	国际数学家大会颁发2006年菲尔兹奖给佩雷尔曼,但他拒绝受奖.

六、后记

庞加莱猜想的解决,以及佩雷尔曼不同于一般人的抉择,引起媒体广泛的报导,但在媒体的操作下,也引起许多扭曲及风波纷扰.例如一再呈现在公众面前的数学家形象,均是如电影“美丽心灵”中的纳什(John Nash)或现在的佩雷尔曼等较殊异的特例.事实上,数学家不是一个怪异族群,大部分都和其他人一样,拥有正常的人际关系及家庭生活,在自己工作岗位上兢兢业业地努力,才能在研究上有些许的突破.2006 年菲尔兹奖得主陶哲轩(Terence Tao)等人,也都和佩雷尔曼一样杰出,我们可以预见他们将继续引领数学界数十年的发展,然而这几位聪明的年轻数学家,显然都被媒体忽略了.

佩雷尔曼对数学研究的全然投入与执著令人景仰及动容,他的物质需求很低,生活简单得近似苦行僧,他认为研究的努力在问题被解决的那一刻,即已得到馈赏及回报,并不需要其他桂冠.他的朋友,也是来自前苏联的史丹佛大学教授伊利沙柏格(Yakov Eliashberg),臆测当初佩雷尔曼会离开美国回到圣彼得堡的原因,正是为了专心解决庞加莱猜想.

佩雷尔曼的特立独行其来有自,在 1990 年代他就曾拒绝接受欧洲数学学会颁给他的奖项,他认为他们无法评断他的工作;1995 年一些美国知名大学邀请他申请教职时,他也拒绝寄履历表及推荐信,因为如果他们了解他的工作,就不需要这些资料,如果需要这些资料,则显然并不了解他的工作.佩雷尔曼希望能完全不受外界干扰地做他自己想做的研究,外界及媒体过多的关注及瞩目,或许才是佩雷尔曼再度封闭自己的主要原因.

除了佩雷尔曼外,其他投入这项研究的数学家同样令人佩服.汉米尔顿投入瑞奇流的研究近 30 年,他长期的坚持是这个领域能开花结果的主要原因,期间他还曾经因为发表的论文数量不够多,而没得到美国国科会的资助,但他并不因此而放弃.其他如曹怀东、周培能,也都是长期耕耘这个领域的专家.

另一位补充庞加莱猜想证明的中国数学家朱熹平,并不是在中国最受重视的明星学者,他在转换领域时,曾经四五年没发表什么文章,但他不急功近利和一味追逐论文的数量,而是踏实地持续努力,坚持做重要的好问题.不管有什么风波及纷扰,他和曹怀东的工作毫无疑问是世界最顶尖的,也是汉米尔顿眼中非常杰出的研究学者.

另外两组数学家克莱那和洛特以及摩根和田刚,他们愿意接受委托,投注这么多的心力,检查及验证佩雷尔曼的论文,解答众人心中的疑惑,让数学界能在这个基础上继续前进,他们无私的奉献同样令人由衷钦佩.

2006 年 7 月丘成桐院士曾在台湾召开记者会,说明曹怀东和朱熹平的工作.他的目的有两个,一个是鼓励在台湾从事研究的数学家,即使不是身处世界





的学术中心，也依然能够像朱熹平一样做出世界一流的工作.另一个目的，则是要提醒台湾的官员及学术界，不要一味追逐 SCI 论文数，研究工作的质远比量还要重要.然而可惜的是，在媒体的报导下，大家的焦点均落在克莱数学研究所的百万奖金上，这实在是扭曲了丘院士的本意！

2 格里戈里·佩雷尔曼

佩雷尔曼是一位俄罗斯数学家.他是一位 Ricci 流的专家，他可能证明了数学中一个重要的未解决的问题：庞加莱猜想.

佩雷尔曼在著名的圣彼得堡学院就学，他的专业深造是高等数学和物理程序.1982 年他参加国际数学奥林匹克竞赛并获金牌.他在圣彼得堡国家大学数学和机械系获博士学位，此后他在圣彼得堡俄罗斯科学院的 Steklov 数学研究所工作.

20 世纪 80 年代末和 90 年代初佩雷尔曼到美国多所大学工作，他于1995 年或 1996 年回到俄罗斯并重新在 Steklov 研究所工作.到 2002 年秋为止他以他在比较几何方面的工作而知名.在这个方面他获得了一些可观的结果.

2002 年 11 月他在预印本文献库发表了一篇文章，这是一系列文章的第一篇.这些文章似乎说明佩雷尔曼证明了几何化猜想，这个猜想的一个特殊情况就是庞加莱猜想.许多人认为，法国数学家亨利·庞加莱于 1904 年提出的庞加莱猜想是拓扑学最著名的未解决的问题.许多数学家徒劳地试图证明这个问题，克莱数学研究所为它的解决出赏金 100 万美元.

佩雷尔曼的解题方案在于使用 Ricci 流来改变理查得·哈密尔顿的几何化方法.与直接的拓扑学方案相比这个方案似乎更可行.

到 2004 年 9 月为止，数学界仍在检查佩雷尔曼的证明，他本人在一些知名的大学里讲课来解释他在预印本文献库发表的文章中的证明.

有人怀疑假如他的证明被确认是正确的，他是否会获得那 100 万美元奖金.20 世纪 90 年代初他拒绝接受欧洲数学协会的一个奖金.有人说他“非常不物质主义”，至今为止他也不打算将他的证明发表在任何由同行评价的数学杂志上.而在一份同行评价的杂志上发表其证明是获得该奖金的条件之一.另一方面，至今为止其他数学家对他在预印本文献库上发表的文章的检查已经远超过了杂志中的同行评价，该奖金的委员会表示在这种情况下它可能修改获得奖金的条件，对预印本文献库来说，这是一个非常重要的步骤，因为它可能将预印本文献库提高到与传统出版物相同的地位.

美国《科学》杂志公布了该刊评选出的 2006 年度十大科学进展，其中科学家证明庞加莱猜想被列为头号科学进展.

《科学》杂志说，科学家们在2006年完成了“数学史上的一个重要章节”，这个“有关三维空间抽象形状”的问题终于被解决.庞加莱猜想属于数学中的拓扑学分支，1904年由法国数学家庞加莱提出，即如果一个封闭空间中所有的封闭曲线都可以收缩成一点，那么这个空间一定是三维圆球.百余年来，数学家们为证明这一猜想付出了艰辛的努力.

被称为数学隐士的俄罗斯数学家佩雷尔曼在证明庞加莱猜想过程中发挥了最为重要的作用，后来三个独立的小组又进一步填补了佩雷尔曼证明中缺失的关键细节，百年难题终获破解.《科学》杂志称，科学家们已经达成共识，认为这一猜想已经被证明.

2006年菲尔兹奖得主之一、俄罗斯数学家格里戈里·佩雷尔曼是一个神秘人物.自从在因特网上发表3篇庞加莱猜想的关键论文之后不久，佩雷尔曼就不再露面，甚至连菲尔兹奖可能也无法把他吸引出来.

实际上，佩雷尔曼在他的学术生涯中曾多次拒绝荣誉或奖项.1995年，他拒绝斯坦福大学等一批美国著名学府的邀请；1996年，他拒绝接受欧洲数学学会颁发的杰出青年数学家奖.

自2002年11月起，佩雷尔曼先后把3页简短的论文贴到互联网上，宣布他解决了七大数学难题之一的“庞加莱猜想”.4个月后佩雷尔曼又在网上公布了第二份报告，介绍了证明的更多细节.同时他也通过电子邮件与该领域的少数专家进行交流.

但是，旋风般地在美国进行了巡回演讲并取得成功后，佩雷尔曼于2003年春从人们的视线里消失，留下全世界数学家费尽心力一行一行地梳理他的论文，填补细节并寻找论文在逻辑上的漏洞.

“如果有人对我解决问题的方式感兴趣，它就在那.”佩雷尔曼说，“我公布了所有的计算.这是我能提供给公众的.”

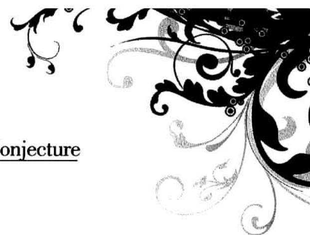
美国《纽约时报》近日的一篇报道，开头就是“佩雷尔曼，你在哪里?”据说，美国数学界对这位天才极其佩服，但他拒绝了斯坦福大学、普林斯顿高等研究院等著名学府的聘请，而宁可“在圣彼得堡附近的森林里找蘑菇”.

美国数学家说，不修边幅的佩雷尔曼“友善而害羞，对一切物质财富不感兴趣”，他“似乎不是生活在这个世界的人”.纽约州立大学数学家迈克尔·安德森说，“佩雷尔曼来过了，解决了问题，其他的一切对于他都是肤浅的.”

佩雷尔曼的拒领大奖和深居简出引起了人们的猜疑，一些人认为他也许正在攻克其他什么学术难题.

佩雷尔曼曾在圣彼得堡接受采访时表示，自己根本不值得人们如此关注.他说:“我不认为自己说的话能引起公众的兴趣.我不说，是因为我重视隐私，而不是我隐藏了自己正在做的事情.没有什么所谓的顶级计划正在进行.我只是





认为公众对我根本没有兴趣."

据俄罗斯数学家诺维科夫介绍,佩雷尔曼早在 1992 年就已开始了"庞加莱猜想"的证明,在此后漫长的过程中,佩雷尔曼除了这个猜想心中已经没有其他.

不管清贫与否,佩雷尔曼生活简单是无疑的.据他的邻居们说,佩雷尔曼一直过着隐居的生活.除了会定时光顾离家不远的一个副食商店外,他基本不离开自己的家.

据该副食商店商品检验员奥丽加·明茨和塔季扬娜·波里亚科娃介绍,许多年来,佩雷尔曼买的东西基本没有改变过:一个黑面包,一包通心粉,比菲多克牌和比菲来弗牌酸奶.水果部那边他几乎都不过去,进口苹果和橙子他似乎买不起.他也不买酒水和其他多余的东西.总之,"只买那些很便宜又好做的简单食品."

是否因为像外界所传言"因路费问题"佩雷尔曼才未去领奖时,基斯里亚科夫否认了这种可能,因为国际数学家大会组委会会为他支付一切相关费用.另外,拒绝领取奖章和接受奖金是两码事.据悉,即使佩雷尔曼没有到颁奖现场,美国克莱数学研究所也须在 2 年内就该问题成立专门委员会,来最终决定是否给他颁发这笔奖金.

"一身黑色的衣服,长长的头发,长长的指甲,一成不变的食品,总是在同一个时间来商店……"

如果这笔奖金最终送达佩雷尔曼手里,也许他应该考虑整理一下服饰,顺便购物时多些花样,以便奥丽加·明茨和塔季扬娜·波里亚科娃不再觉得他像上面那样,如同幽灵.

3 朱熹平

朱熹平,男,1982 年本科毕业于中山大学数学系,1984 年在中山大学数学系取得硕士学位,1989 年在中国科学院武汉数学物理研究所取得博士学位.现任中山大学数学系教授、博士生导师,数学与计算科学学院院长,兼任广东省数学学会理事长,中国科学院晨兴数学中心学术委员会委员,浙江大学数学科学研究中心顾问.朱熹平于 1991 年获中国科学院自然科学二等奖,于 1997 年入选教育部"跨世纪人才培养计划",1998 年获国家杰出青年基金,列入 1999 年度人事部"百千万人才工程"第一、第二层次人选,并于 2001 年被聘为教育部"长江学者奖励计划"特聘教授,2004 年获得全球华人数学家大会颁发的晨兴数学银奖.

4 曹怀东

曹怀东:Lehigh 大学数学系讲座教授.

曹怀东教授是美国里海(Lehigh)大学数学系的 A. Everett Pitcher 讲座教授.

曹怀东 1981 年本科毕业于清华大学,1986 年在 Princeton 大学数学系获得博士学位,师从著名数学家丘成桐.曹怀东主要从事的研究领域是微分几何学,涉及 Kahler-Ricci 流、数学物理等众多方面.

曹怀东曾获得 Alfred P. Sloan 基础研究奖金(1991 ~ 1993),John Simon Guggenheim 基金会奖金(2004)等,是国际著名数学杂志 *Journal of Differential Geometry* 的执行主编.

曹怀东以及另一位 Yau 的学生周培能(Bennett Chow)在 Ricci 流的研究中做出了许多重要的工作,受到 Ricci 流理论的创始者美国科学院院士理查德·哈密尔顿的高度评价.

令人特别佩服的是,曹怀东在国外的前 4 篇文章,分别发表在 1985,1986,1990,1992 年的 *Inventiones Mathematicae* 上.一般认为,目前最顶尖的数学综合性杂志(不包括 *JDG*,*Topology* 这样的专业性顶尖杂志)是 *Inventiones Mathematicae*,*Annals of Mathematics*,*Acta Mathematica* 以及 *Jounal of AMSo*.国内的教授如果能有一篇论文发在上述杂志上,基本上评博导是没有问题的.

哈密尔顿从 Eells-Sampson 的调和映照热流的工作受到启发,在 1982 年的文章中首先引入 Ricci 流的概念,就是从一个给定的初始黎曼度量出发,依照其 Ricci 曲率的变化,通过解一个抛物型的发展方程,得到一个单参数族的黎曼度量,最后希望证明当参数趋于无穷时,收敛到一个常曲率的度量.

哈密尔顿在 1982 年证明了三维闭流形上当初始度量具有正的 Ricci 曲率时,Ricci 流方程的解在任意时刻都存在,并且收敛到一个正的常截曲率度量.大家回忆一下庞加莱猜想,闭的单连通 3 维拓扑流形若和 3 维球面有相同的同调群(或等价地说,和 3 维球面同伦),则其实同胚于 3 维球面.而哈密尔顿的这个曾轰动一时的发现使得我们只要证明任何一个闭的 3 维同伦球面上都存在正的 Ricci 曲率度量,就证明了庞加莱猜想.这个猜想是庞加莱在 1904 年提出来的.这是一个让无数拓扑学家迷恋的难题,错误的证明层出不穷,对它的研究极大地推动了 3 维拓扑学的发展.人们把它推广到高维情形的广义庞加莱猜想,是说 n 维闭流形若同伦于 n 维球面,则其实同胚于 n 维球面.$n>4$ 被 Smale 证明,$n=4$ 被 Freedman 证明.Thurston 从更高的观点考虑庞加莱猜想,提出了椭圆化猜想,是说每个具有有限基本群的闭的 3 维流形上,都存在正的常曲率度量.由于具有常正曲率的 3 维流形都以 3 维球面为它的万有覆盖.





所以从这个椭圆化猜想就可以推出庞加莱猜想.椭圆化猜想是更广的Thurston几何化猜想的一个特例(对应球面几何情形).Smale,Freedman,Thurston都是由于以上这些所提到的工作,获得了菲尔兹奖.可见这个问题的重要.

由于一般来说,Ricci流方程的解在有限时间后会出现奇点,所以哈密尔顿又引入了对拓扑流形进行外科手术的技巧,和分析工具结合使用,得到了一大批激动人心的结果,建立起了一整套的证明Thurston几何化猜想的框架.他的工作,使得人们相信,只要沿着Ricci流的方向走下去,迟早能解决这个让拓扑学家无能为力的难题.最近,俄国数学家佩雷尔曼宣称已经完全证明了哈密尔顿框架里的关键步骤,从而也彻底解决了Thurston的几何化猜想.他的工作虽然还在审查中,但从目前得到的信息来看,是非常乐观的.可以确定的是,佩雷尔曼的工作极大地推动了Ricci流的发展,促进了分析学和拓扑学的融合.

5 丘成桐

想起数学家,我们自然会想起国际数学界的泰斗、2002年国际数学家大会名誉主席陈省身先生,他一生不仅在几何学上取得巨大的成就,还为祖国培养了很多世界级的数学家,吴文俊是其中一位.而我们今天要介绍的这位潮汕人则是他的另一位高徒,他于1983年获得菲尔兹奖,这是世界数学领域的诺贝尔奖,直到今天,他还是华人中唯一的获奖者,这个殊荣就连陈省身先生都没有获得过,他就是丘成桐.

丘成桐26岁便登峰造极,成就了一生的辉煌.他最重要的工作在26岁之前就完成了.几十年来他对数学永不停歇的追求,让他的名字在几何、拓扑和理论物理学界几乎无处不在.国外一个很有名的数学家劳森说,他和丘教授合作的时候,丘教授还没毕业,但当时他就感觉到这个年轻人的名字将会在几何领域里处处稠密,但他没有想到如今这个名字在理论物理界也是处处稠密.大家应该有稠密这个概念,就是说总是能听到看到这个名字,比如说开超弦大会,卡拉比-丘流形,几乎每一个演讲里都会提到这个名词.

几十年来数学就是他的生活,他的娱乐,他的生命.跟他在一起很累,因为你要不停地跟着他的脑子想数学.与他相处,开始时或许你不会觉得他有任何过人之处,大家看他也许普普通通,甚至连普通话都讲得不是很清楚.但在讨论问题的时候,他直达难题本质的气魄,只能是天生才有,这是天生的大气魄.因为解决问题时,一般人有时总想要一点小技巧,可他碰到难题就是硬要把它砸开.有人称他是数学界的"凯撒大帝".

他似乎是为数学而生,数学因他变得更精彩,他做人、做事、做学问,都有一种大开大合、征服一切的勇气,在数学内外他都有着无与伦比的影响力和感染力.

5.1 人物背景

丘成桐1949年出生于广东汕头，后全家定居香港．父亲曾在香港香江学院及香港中文大学的前身崇基学院任教．父教母慈，童年的丘成桐无忧无虑，成绩优异．但在他14岁那年，父亲突然辞世，一家人顿时失去经济来源．尽管丘成桐不得不一边打工一边学习，却仍然以优异成绩考入香港中文大学数学系．

他中学的时候逃学一年，曾经成绩很差，差一点落榜．19岁的时候来到美国伯克利，“21岁毕业时就注定要改变数学的面貌”．这是几年前加州大学洛杉矶分校希望把丘成桐聘请过来的时候，系里讨论时一个年纪很大的几何学家引用陈省身先生说的一句话．他10年之后成为数学界的一代天骄．从他入学伯克利到在世界数学家大会做一小时报告还不到10年，当年他只有28岁，也是在那一年，陈景润先生被邀请作45分钟的报告．这期间他证明了卡拉比猜想、正质量猜想，开创了一个崭新的领域：几何分析．

1981年，他32岁时，获得了美国数学会的维布伦(Veblen)奖——这是世界微分几何界的最高奖项之一；1983年，他被授予菲尔兹(Fields)奖章——这是世界数学界的最高荣誉；1994年，他又荣获了克劳福(Crawford)奖．

除此之外，他还获得过美国国家科学奖章和加利福尼亚州最优秀的科学家的称号，是美国科学院院士、哈佛大学名誉博士、中国科学院外籍院士、香港中文大学名誉博士……

大学期间，他以三年时间修完全部必修课程，还阅读了大量课外资料．他的突出成绩和钻研精神为当时的美籍教授萨拉夫所赏识，萨拉夫力荐他到美国加利福尼亚大学伯克利分校攻读博士研究生．20世纪70年代左右的伯克利分校是世界微分几何的中心，云集了许多优秀的几何学家和年轻学者．在这里，丘成桐得到IBM奖学金，并师从著名微分几何学家陈省身．

5.2 过程

数学是奇妙的，也是艰涩的．即使是立志在数学领域建功立业的年轻学生，能坚持到最后并出成果的，也是寥若晨星．丘成桐正可谓是这样一颗“晨星”．常常有这样的情景——偌大的教室中，听课的学生越来越少，最后竟然只剩下教授一人面对讲台下唯一的学生悉心教诲．这唯一的学生，就是丘成桐．到伯克利分校学习一年后，丘成桐便完成了他的博士论文，文中巧妙地解决了当时十分著名的“沃尔夫猜测”．他对这个问题的巧妙解决，使当时的世界数学界意识到一个数学新星的出现．

丘成桐取得博士学位后，在应邀前往普林斯顿高等研究院访问的一年中，他结识了许多年轻的世界一流数学家，完成了两篇论文．1972年秋，年仅23岁的丘成桐应邀来到纽约大学石溪分校担任副教授，又完成了几篇论文．在1973





年美国数学会举行的微分几何大会上，丘成桐做了三个学术报告，以卓越的能力和杰出的贡献，向数学界显示了自己在微分几何领域的领先水平．这一年是丘成桐数学事业上十分重要的一年，他完成了题为“完备黎曼流形上调和函数”的著名论文，用他自己的话说，这篇文章是他数学生涯的转折点．实际上，该文奠定了他应用分析方法的基本思想和技巧．

丘成桐最重要、最有影响的工作是对“卡拉比猜想”的证明．他是在1976年底用强有力的偏微分方程估计解决了这一问题的．在解决“卡拉比猜想”的同时，他还证明了负定第一陈类的紧克勒流形上克勒－爱因斯坦度量的存在性．

1976年，丘成桐被提升为斯坦福大学数学教授．1978年，他应邀在芬兰举行的世界数学大会上做题为“微分几何中偏微分方程应用”的学术报告．这一报告代表了20世纪80年代前后微分几何的研究方向、方法及其主流．这之后，他又解决了“正质量猜测”等一系列数学领域难题．

丘成桐的研究工作深刻又广泛，涉及微分几何的各个方面，成果累累．1989年，美国数学会在洛杉矶举行微分几何大会，丘成桐作为世界微分几何的新一代领导人出任大会主席．

5.3　精神

命运是公平的，奖章、荣誉授予了那个在教室中坚持到最后的人．这，并没有让丘成桐止步不前，他继续进行着大量繁杂的研究工作，并不断取得成就．

坚韧、坚持、锲而不舍，这就是丘成桐的精神．当然，也不是每个有着这样精神的人都能取得丘成桐一样的成就．数学需要勤奋，更需要天才．正如著名数学家尼伦伯格所说，丘成桐“不仅具备几何学家的直观能力，而且兼有分析家的才能”．著名数学家郑绍远先生回忆说，对于许多艰深的数学问题，丘成桐已思考了近20年，虽然仍未解决，他还是没有轻易放弃．

丘成桐对中国的数学事业一直非常关心．从1984年起，他先后招收了十几名来自中国的博士研究生，要为中国培养微分几何方面的人才．他的做法是，不仅要教给学生一些特殊的技巧，更重要的是教会他们如何领会数学的精辟之处．他的学生田刚，也于1996年获得了维布伦奖，被公认为世界最杰出的微分几何学家之一．

数学是奇妙的，只有锲而不舍才能探求其中真谛．对于丘成桐这样的数学家来说，这种探求不但是人生的意义，也是人生的乐趣．

丘成桐不是一个完人，但绝对是一个伟大的数学家．你可以不喜欢这个人，但你不可能不喜欢他的数学，他证明了许多妙不可言的定理．大家如果学数学，读到研究生的话你就会知道他的定理非常美妙，他对卡拉比猜想的证明毫无疑问是数学中最深刻的理论之一，尤其是在超弦理论中应用之广不可思议．

几十年来，丘先生不仅时刻把握着数学与物理跳动的脉搏，引导着世界数

学发展的潮流,还一直怀着一颗赤子之心,关心和帮助着中国数学的进步.他培养了众多的华人数学家.他的学生和博士后在国外各个重要的大学里几乎都有.

5.4 尾声

他创建了4个数学中心,目的就是为培养中国的一代年轻人,他把它作为自己一生事业的重点.现在他每天忙着待人接物,其实很多目的是捐款,希望能够为国家省一些钱,用富商们的捐款来培养中国的年轻人.丘教授的目标就是希望过5年、10年能够再培养出几个朱熹平.他创办了世界华人数学家大会,让无数海内外数学家受益.

一些仰望他的无名学子,有幸成为他的弟子,又成为他的合作者和朋友,相互的交往让这些学子对他的尊敬和爱戴与日俱增.做学生时看到波士顿的科学博物馆里数学馆的墙壁上镌刻着几十个当代伟大数学家的名字,其中有3个中国人:华罗庚、陈省身、丘成桐.丘教授的名字刻上去的时候才三十几岁.到如今想到几何,想到物理,想到中国的科学,当今世上任何一个数学家或者物理学家都会想到这个名字——丘成桐,一个镌刻在数学史上的中国人的名字.

6 菲尔兹奖

1936年开始颁发的菲尔兹(Fields)奖被称为数学界的诺贝尔奖.

菲尔兹奖是以已故的加拿大数学家、教育家J·C·菲尔兹(Fields)的姓氏命名的,中文全名:约翰·查尔斯·菲尔兹.

菲尔兹奖是最著名的世界性数学奖,由于诺贝尔奖没有数学奖,因此也有人将菲尔兹奖誉为数学中的“诺贝尔奖”.

J·C·菲尔兹1863年5月14日生于加拿大渥太华.他11岁丧父,18岁丧母,家境不算太好.J·C·菲尔兹17岁进入多伦多大学攻读数学,24岁时在美国的约翰·霍普金斯大学获博士学位,26岁任美国阿勒格尼大学教授.1892年他到巴黎、柏林学习和工作,1902年回国后执教于多伦多大学.J·C·菲尔兹于1907年当选为加拿大皇家学会会员.他还被选为英国皇家学会、苏联科学院等许多科学团体的成员.

J·C·菲尔兹强烈地主张数学发展应是国际性的,他对于数学的国际交流,对于促进北美洲数学的发展都抱有独特的见解并满腔热情地作出了很大的贡献.为了使北美洲数学迅速发展并赶上欧洲,他第一个在加拿大推进研究生教育,也是他全力筹备并主持了1924年在多伦多召开的国际数学家大会(这是在欧洲之外召开的第一次国际数学家大会).正是这次大会使他过分劳累,从此健





康状况再也没有好转,但这次大会对于促进北美的数学发展和数学家之间的国际交流,确实产生了深远的影响.当他得知这次大会的经费有结余时,他就萌发了把它作为基金设立一个国际数学奖的念头.他为此积极奔走于欧美各国谋求广泛支持,并打算于1932年在苏黎世召开的第九次国际数学家大会上亲自提出建议.但不幸的是未等到大会开幕他就去世了.J·C·菲尔兹在去世前立下了遗嘱,他把自己留下的遗产加到上述剩余经费中,由多伦多大学数学系转交给第九次国际数学家大会,大会立即接受了这一建议.

J·C·菲尔兹本来要求奖金不要以个人、国家或机构来命名,而用"国际奖金"的名义.但是,参加国际数学家大会的数学家们为了赞许和缅怀J·C·菲尔兹的远见卓识、组织才能和他为促进数学事业的国际交流所表现出的无私奉献的伟大精神,一致同意将该奖命名为菲尔兹奖.

第一次菲尔兹奖颁发于1936年,当时并没有在世界上引起多大注意.连许多数学专业的大学生也未必知道这个奖,科学杂志也不报道获奖者及其业绩.然而30年以后的情况就完全不一样了.每次国际数学家大会的召开,从国际上权威性的数学杂志到一般性的数学刊物,都争相报导获奖人物.菲尔兹奖的声誉不断提高,终于被人们确认:对于青年人来说,菲尔兹奖是国际上最高的数学奖.

菲尔兹奖的一个最大特点是奖励年轻人,只授予40岁以下的数学家(这一点在刚开始时似乎只是个不成文的规定,后来则作出了明文规定),即授予那些能对未来数学发展起到重大作用的人.

菲尔兹奖的奖品是一枚金质奖章和1 500美元的奖金,奖章的正面是阿基米德的浮雕头像.就奖金数目来说与诺贝尔奖金相比可以说是微不足道,但为什么在人们的心目中,它的地位竟如此崇高呢? 主要原因有三:第一,它是由数学界的国际权威学术团体——国际数学联合会主持,从全世界的第一流青年数学家中评定、遴选出来的;第二,它是在每隔四年才召开一次的国际数学家大会上隆重颁发的,且每次获奖者仅2~4名(一般只有2名),因此获奖的机会比诺贝尔奖还要少;第三,也是最根本的一条是由于得奖人的出色才干,赢得了国际社会的声誉.正如20世纪著名数学家C·H·H·外尔对1954年两位获奖者的评介:他们"所达到的高度是自己未曾想到的","自己从未见过这样的明星在数学天空中灿烂升起","数学界为你们二位所做的工作感到骄傲".从而证明了菲尔兹奖对青年数学家来说是世界上最高的国际数学奖.

菲尔兹奖的授奖仪式,都在每次国际数学家大会开幕式上隆重举行,先由执委会主席(即评委会主席)宣布获奖名单,接着由东道国的重要人物(当地市长、所在国科学院院长甚至国王、总统)或评委会主席或众望所归的著名数学家授予奖章和奖金.最后由一些权威数学家分别、逐一简要评介得奖人的主要数学成就.

最后一位通才——庞加莱[1]

埃里克·坦普尔·贝尔(Eric Temple Bell)1883年出生于苏格兰的阿伯丁.早年就学于英格兰.1902年到美国,进入斯坦福大学学习,1904年取得文学学士学位.1908年在华盛顿大学做研究生,并从事教学,1909年获该校文学硕士学位.1911年进入哥伦比亚大学,1912年获该校哲学博士学位.此后回华盛顿大学任数学讲师.1921年晋升为教授.1924~1928年夏任教于芝加哥大学.1926年上半年任教于哈佛大学,随之受聘为加州理工学院的数学教授.贝尔是美国国家科学院院士.曾任美国数学协会主席.美国数学学会和美国科学促进会副主席.贝尔于1960年12月逝世.

第一章

一个名副其实的科学家,尤其是一个数学家,在他的工作中会感受到与一个艺术家同样的印象;他的愉快也同样巨大,并具有同样的性质.

——亨利·庞加莱

星占学家威廉·利利(William Lilly, 1602—1681)在《他的生活和时代的历史》中,对于对数发明者、默契斯登堡的约翰·纳皮尔(John Napier,1550—1617)和计算出第一个常用对数表的、伦敦格雷沙姆学院的亨利·布里格斯(1561—1631)的会见,作了有趣的——即便是难以置信的——记录.一个叫约翰·马尔(John Marr)的"杰出的数学家和几何学家","在布里格斯先生之前到达苏格兰,目的是要在这两个如此博学的人会面时在场.布里格斯先生指定了在爱丁堡会面的日期,但是未能如期前往,纳皮尔勋爵怀疑他不会去了.一天,碰巧马尔和纳皮尔勋爵正讲到布里格斯先生:'噢,约翰(默契斯登堡的约翰说),布里格斯先生不会来了.'正在这时,有人敲门;约翰·马尔急忙下楼,他很满意地证实来人正是布里格斯先生.他把布里格斯先生领上楼,带进勋爵的房间,在开口说话之前,两人互相赞赏地注视着,几乎过了一刻钟."

① 原载自:E·T·贝尔著.从芝诺到庞加莱——数学大师.徐源,译.宗蜀碧,校.上海:上海科技教育出版社,2004:631-663.





西尔维斯特回想起这个传说，讲述了他自己怎样仿效布里格斯那令人目瞪口呆的表示赞赏的世界纪录，那是在 1885 年，他访问一位有许多令人惊奇地成熟而不可思议的创见性文章的作者时的事情，这些有关分析学的一个新分支的文章，从 19 世纪 80 年代初开始，就使各种数学杂志的编辑们应接不暇了.

"当我最近在盖 - 吕萨街庞加莱那通风的休息处拜访他时，"西尔维斯特承认，"我完全进入了布里格斯面晤纳皮尔时的感情……在那个被抑制的智慧的伟大蓄积者面前，我的舌头一下子失去了功能，直到我用了一些时间(可能有两三分钟)仔细端详和承受了可谓他思想的外部形式的年轻面貌时，我才发现自己能够开口说话了."

西尔维斯特在另一处记述他辛辛苦苦地爬上通往庞加莱那"通风的休息处"的三层窄楼梯以后，停下来擦着他那硕大的秃头时，看到的不过是一个孩子，当时他惊奇得不知所措，这个"如此美貌，如此年轻"的孩子，竟然是那些洪水般涌来的文章的作者，他预告了柯西的一个后继者的到来.

第二件轶事，多少可以说明那些能够欣赏庞加莱的工作范围的人，对他的工作所持的尊敬程度.当一位爱国的英国将官在(第一次)世界大战的狂热民族主义时期——那时学术界所有的爱国者都把抬高他们风雅的盟国，贬低他们粗俗的敌人当做一种义务——问伯特兰·罗素，法国现代产生的最伟大的人物是谁，罗素立刻答道："庞加莱.""什么？那个家伙?"他那位穿军服的交谈者喊道，认为罗素指的是法兰西共和国的总统雷蒙·庞加莱(Raymond Poincaré)."噢，"当罗素明白了那个人惊愕的原因时，他解释说："我想到的是雷蒙的堂兄弟，亨利·庞加莱."

庞加莱是最后一个实际上以全部数学——包括纯粹数学和应用数学两方面——作为他的研究领域的人.人们一般认为，在今天起步的任何人都不可能全面地理解数学的四个主要部分——算术、代数、几何、分析——中的两个以上，更不用说做出高质量的创造性工作了，对天文学和数理物理学就更谈不上了.但是，甚至在 19 世纪 80 年代，在庞加莱的伟大事业开始时，人们就普遍认为高斯是最后一个数学通才(也有人认为是希尔伯特)，所以不能证明不可能有某个未来的庞加莱再次包罗整个领域.

至于数学的演化，它既膨胀也收缩，多少有点像一个勒梅特的宇宙模型.目前，这个状况是爆炸式地膨胀的，任何人要想熟悉自 1900 年以来倾注于世的全部刚刚开始的数学，是完全不可能的.但是在一些重要的部分，一种最受欢迎的收缩倾向是清楚明显的.例如，在代数中就是如此，公理法的全面采用，立刻使得这个学科更抽象、更一般、更连贯.现代着手的某些工作揭开了一些意想不到的类似情形.可以想象，下一代的代数学家将无须知道许多现在被视为有价值的东西，因为这些特殊的、困难的东西，有许多将被归入范围更广阔的、较为简

单的一般原理.当相对论把以太的复杂数学束之高阁时,在经典的数理物理学中就发生了这类事情.

在膨胀中的这种收缩的另一个例子,是较之向量分析的各种各样用途来说,张量分析的应用有了迅速增长.这样的推广和浓缩一开始往往很难为年纪大的人所掌握,并且常常要经过严峻的斗争才能存在下来,但是最后人们通常认识到,一般方法比为特殊问题设计的各种各样巧妙的花样在本质上更简单、更容易掌握.当数学家们断言诸如张量分析这样的东西是容易的——至少与它之前的一些算法相比较时,他们并不是要显得优越或神秘,而是在陈述一条每个学生都自己证实的、有价值的真理.这种范围广大的普遍性的特性,是庞加莱大量著作的一个卓著特点.

如果抽象性和普遍性具有所指出的那种明显的优点,那么它们有时候对那些关心细节的人来说有严重缺点也是事实.一个工作中的物理学家知道在其工作中出现的特殊微分方程是可解的,因为当他和数学家都不能完成可用于特殊问题的某种数值解所要求的赫拉克勒斯①式的劳动时,某个纯粹数学家证明了它可解,可是这对他有什么直接用处呢?

在庞加莱做出他的一些最有独创性的工作的领域中,举一个例子:考虑一个均匀的、不可压缩的流体,被它的粒子的引力作用聚合在一起,围绕一个轴旋转.在什么条件下运动是稳定的? 这样一个稳定的旋转流体的形状是什么? 麦克劳林、雅可比和其他一些人证明了一些椭球是稳定的;庞加莱采用了比他的前辈们更直观,而且"不那么数学"的方法,他一度认为他确定了一个梨形体的稳定性的判据.但是他错了,他的方法不适于数值计算,后来的工作者,包括著名的查理·达尔文的儿子 G·H·达尔文,在得出确定的结论以前,没有被必须从道路上清除的极其可怕的代数和算术的丛林阻挡住,他们得到了一个肯定的解.②

如果数学家们的发现能以某种可使用计算机器的形式提交给一位对双星演化感兴趣的人,那么他就会感觉较为舒适.自从克罗内克的"没有构造就没有存在"(这一名言的另一版本是:"整数是上帝创造的,其余都是人造的")的法令以来,一些纯数学家本人对于非构造性的存在定理,也不像他们在庞加莱的时代那样热心了.庞加莱蔑视数学的使用者们所要求的在进行工作以前必须具备的那些细节,这是促进其工作的普遍性的最重要的原因之一.另一个原因是他非同寻常地全面掌握了单复变量函数理论的全部方法.在这方面他是无与伦比

---

① 赫拉克勒斯为希腊神话中主神宙斯之子,力大无穷,曾完成 12 项英雄业绩.——译注者

② 这个著名的"梨形体"问题,在天体演化论中有相当大的重要性.1905 年由李雅普诺夫(Liapounoff)明白无误地解决了,他的结论在 1915 年由金斯爵士(Sir James Jeans)进一步证实:他们发现运动是不稳定的.极少有人有勇气去检查那些计算.1915 年以后,李雅普诺夫的一位同胞莱昂·李希滕施泰因(Lam Lichtenstein)对旋转流体质量问题作了全面考虑.这个问题似乎会招致不幸:两个李都死于非命.





的.可以注意到,庞加莱把他多方面的才能极好地用来揭示以前未曾猜想到的、数学的相距甚远的分支——例如(连续)群和线性代数——之间的联系.

为了全面起见,我们在讲到庞加莱的生平以前,必须回想他的观点的另一个特点:很少有几个数学家具有庞加莱那样宽广的哲学观点,没有一个数学家在清晰地讲解的天赋方面超过他.也许他总是对科学和数学的哲学含意深感兴趣,但是只是在1902年,当他作为一个伟大的专业数学家的地位,已经无可动摇地确立起来时,他才把普及数学从一种附带的兴趣,转到可以称为数学普及的呼吁,真诚热心地投身于和非专业的人们一起,尽情分享他的学科的意义及其对人类的重要性.在这里,他对普遍胜于对特殊的喜爱,帮助他深入浅出地向知识的门外汉们,讲述数学中比技术更重要的东西.二三十年前,可以在巴黎的公园和咖啡馆里,看到工人和女店员们热心地读着庞加莱的这一部或那一部杰作的印刷简陋、封面纸张低劣的通俗本.这些著作的较好的版本,也可以在有专业学问的人的书桌上找到——经过反复翻动,明显是读过的.这些书被译成英文、德文、西班牙文、匈牙利文、瑞典文和日文出版.庞加莱对所有的人用他们能识别的语言,讲着数学和科学的通用语言.他的风格,他自己特有的风格,在翻译中失去了很多.

由于通俗著作的文学成就,庞加莱被授予法国作家所能得到的最高荣誉,成为法兰西学院文学部院士.心怀妒忌的小说家们有几分恶意地说,庞加莱获得这个对科学界人士独特的荣誉,是因为(文)学院的一个职能是不断地汇编一部权威性的法语字典.通才庞加莱显然是在诗人和语言学家们告诉世界什么是自守函数的努力中给他们帮了忙.通过对庞加莱的著作进行研究,在此基础上确立的公正看法,认为这个数学家正该得到他所得到的荣誉.

与庞加莱对数学的哲学思想感兴趣密切相关的,是他对数学创造心理的专注.数学家们是怎样做出他们的发现的呢?后面庞加莱会在他写的有关个人发现的一个有趣的故事中,告诉我们他对这个秘密的看法.结论好像是数学发现是数学家在这方面的长期艰苦劳动之后,或多或少自己冒出来的.正如在文学中——按照丹特·加布里埃尔·罗赛蒂(Dante Gabriel Rossetti)①的话——一首诗成熟之前,“一定量的基本脑力劳动”是必要的,在数学中也是如此,没有事先的单调辛苦的工作,就没有发现,但是这决不是整个情形.对创造力的全部“解释”如果不能提供某种窍门,使有天赋的人能据以进行创造,那就很值得怀疑.庞加莱到实用心理学中去遨游,和其他一些在同一方向遨游的人一样,没有能带回金羊毛②,但是它至少提出了这样的事情并不是完全神秘的,有一天当人类聪明到能够懂得他们自己的躯体时,金羊毛是可以找到的.

---

① 罗赛蒂(1828—1882),英国画家、诗人,拉斐尔前派创始人之一,祖籍意大利.——译者注
② 见希腊神话中阿尔戈(Argo)英雄随同伊阿宋(Jason)乘舟去海外寻找金羊毛的故事.——译者注

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

庞加莱的智力在父母两方面都很好地得到了遗传.我们只追溯到他的祖父.在1814年的拿破仑战争期间,这位祖父只有20岁,隶属于在圣–康坦的陆军医院.1817年他在鲁昂定居之后,结了婚,生了两个儿子:莱昂·庞加莱,生于1828年,后来成为一名第一流的医生和医师公会的成员;安托万(Antoine),成为公路及桥梁部门的总检察官.莱昂的儿子亨利于1854年4月29日出生在洛林的南希,成为20世纪初的一流数学家;安托万有两个儿子,雷蒙从事法律,在第一次世界大战期间升任法兰西共和国总统,安托万的另一个儿子成了中等教育局局长.庞加莱祖父的一个兄弟跟随拿破仑入侵俄国时失踪了,在莫斯科大败后音信杳无.

从这个杰出的家系表看,人们可能会认为亨利会显现出一些管理才能,但是他没有,只是在他的童年时代随意发明了一些供他的妹妹和小朋友们玩的政治游戏.在这些游戏中,他总是极其公正的,注意让他的每一个玩伴都在"当官"的游戏中扮演他或她的角色.这也许是"从小看大,三岁看老"的一个确证,庞加莱天生不具备懂得最简单的管理原则的能力,而他的堂兄弟雷蒙本能地运用着这些原则.

加斯东·达布(Darboux, Jean-Gaston, 1842—1917),法国数学家.

庞加莱的传记是由他的同胞、现代主要的几何学家之一加斯东·达布在1913年(庞加莱去世后的第二年)非常详细地写成的.有些事本书作者可能有所疏漏,但是好像达布在讲述了庞加莱的母亲"出自默兹区的一个家庭,她的(即这位母亲的)双亲住在阿兰西,她是一个很好的人,很灵敏,很聪明"以后,无意中忘记提到她娘家的姓.难道法国人可能在德国1870年和1914年对法国的文化入侵之后,从他们新近的教师那里接过"三个大写的K"的观点吗?不过,从达布后来讲的一件轶事中,可以知道庞加莱母亲的姓可能是朗努瓦(Lannois).我们知道,这位母亲把她的全部注意力都贡献给她的两个孩子亨利和他妹妹(未提到名字)的教育了.他的妹妹后来成了埃米尔·布特鲁(Emile Boutroux)的妻子和一位数学家(死得很早)的母亲.

庞加莱童年时代智力发展极快,这部分归功于她母亲的不断照料.他很早就学会了说话,但是他在一开始说得很糟,因为他想的比说的要快.从婴儿时起,他的运动的协调性就很差.当他学会写字时,人们发现他左右手都能用,用左手写字或画画同用右手一样糟糕.庞加莱从来没有摆脱这种身体不灵活的毛病.可以回想一下跟这有关的一件趣事,当庞加莱被认为是他那个时代的一流数学家和主要的科学普及者时,他参加了比奈(Binet)测验[①],表现得如此丢脸,以至于要是不把他作为著名的数学家,而是当做孩子来判断的话,他就会被这次测验列为一个低能儿.

① 比奈(1857—1911),法国心理学家,对智力测验做出了重大贡献.——译者注



5 岁时，亨利因患白喉健康受到严重影响，喉咙麻痹达 9 个月之久．这次不幸使他很长时间虚弱而胆小，但是这也使他回到他自己的消遣，因为他不得不避开他那个年龄的孩子们比较粗野的游戏．

他的主要娱乐是阅读，这使他非凡的才能首次显示了出来．一旦读过了一本书——以难以置信的速度——便永远不忘，他总能说出讲到一件特定的事情是在第几页和第几行．他终生保持着这种强有力的记忆力．这种罕见的本领可以称为视觉记忆或空间记忆，欧拉也和庞加莱一样具有这种本领，不过程度差一些．在顷刻记忆——以不可思议的准确性回想起一系列早已过去的事件的能力——方面，他也是强大得不同寻常．然而他还不脸红地说他记忆力“坏”．他可怜的视力也许给他的记忆力增加了第三个奇特之处．大多数数学家看来主要是靠眼睛记住定理和公式，而庞加莱几乎完全是靠耳朵．他在学校学习高等数学时，看不清黑板，就坐在后面听，不记笔记也能很好地跟上并记住——这在他是一种很容易的技术，但对大多数数学家却是不能理解的．然而他一定也有一种“内眼”的生动记忆力，因为他的许多工作就像黎曼的许多工作一样，是伴有敏捷的空间直觉和实际形象的那种类型．他不能灵巧地运用他的手指，这当然妨碍他去做实验．这看来很可惜，因为要是他掌握了实验的艺术，他关于数理物理学的一些工作，可能就与现实更接近了．要是庞加莱在实验科学方面也像他在理论上那么强有力，他可能就会同无与伦比的阿基米德、牛顿和高斯三人一起，成为无与伦比的第四个了．

没有多少大数学家是公众的想象力喜欢把他们描绘成的那种心不在焉的梦想家．庞加莱是一个例外，不过只是在一些较小的事情上，诸如在他的行李中带走了旅馆的亚麻布制品，如台布、床单之类．但是很多决不是心不在焉的人也干同样的事情，人们知道，一些最机灵的家伙把饭店的银餐具偷偷放到衣袋里溜走．

庞加莱的心不在焉的一个方面，类似于一些完全不同的事情(达布没有讲这件事，但是应该讲到它，因为它说明了庞加莱晚年的某些粗暴无礼的行为)．一位著名的数学家从芬兰来到巴黎，跟庞加莱商量一些科学问题，女仆通报后，庞加莱没有离开书房去欢迎这位拜访者，而是继续不停地踱来踱去——因为这是他思考数学时的习惯——足足有三个小时．整个这段时间，那位谦虚的来访者一直安静地坐在隔壁的屋子里，与这位大师只隔着薄薄的门帘．最后，门帘掀开了．庞加莱的水牛般的大脑袋伸进了房间一下．“你大大打扰了我．”那个脑袋突然出现，随即消失了．来访者没有被接见就离开了，这恰恰是“心不在焉”的教授所要的．

庞加莱的小学成绩是优异的．虽然他一开始并没有在数学上显示出任何显著的兴趣．他最早是热爱自然史，他终生都是动物的伟大爱护者．他第一次试用

步枪时,不小心射下了一只他并没有瞄准的鸟.这个不幸深深地影响了他,以至从那以后,什么也不能(除了强迫的军事训练)使他去碰火器.9岁时,他显示出将要成为他的一项主要成就的第一次希望.法文作文的教师宣称,年幼的庞加莱交去的一篇形式和内容都很新颖的短短的习作,是一篇“小小的杰作”,并把它作为他的一件宝贝保留了起来.但是他也劝告这个学生,如果他希望给学校的主考人留个好印象,就得更平常一些——更笨些.

庞加莱不能参加同学们的比较喧闹的游戏,就发明了他自己的游戏.他还成了一个不知疲倦的跳舞迷.由于所有的功课对他来说都像呼吸一样容易,他把他的大部分时间用在玩耍和帮助母亲做家务上.甚至在他一生的这个比较早的阶段,他已显示出了“心不在焉”的一些比较可疑的特点:他常常忘记吃饭,几乎从来记不得他是否吃过早饭.也许他是不愿意像大多数孩子那样吃得过饱.

对数学的热情,在青春期或在那以前不久(当他大约15岁时)抓住了他.从一开始他就显示出一个终生的特点:他的数学是在他不停地踱来踱去时做出来的,而且只是在一切都想好了时才写下来.他工作时,谈话或其他嘈杂的声音从来不会干扰他.后来他写数学论文一挥而就,从来不回头看看他写了些什么,并且在写作时限制他自己只做极少几处删改.凯莱也用这种方法写作,也许欧拉也是这样.庞加莱的一些工作显出匆忙写就的痕迹,他自己说他从来没有在完成一篇文章后,不是为文章的形式,就是为文章的内容懊悔.不止一个写得很好的人有同样的感觉.庞加莱对古典文学研究的鉴赏力,教会了他形式和内容的重要性,他在中学时代古典文学就很好.

1870年庞加莱16岁时,普法战争在法国领土上爆发了.虽然庞加莱太年轻、体质太弱,不能服现役,然而他完全承担了他那一份恐惧,因为他居住的地方南锡,被侵略的浪潮淹没了,这个年轻的孩子便伴随他那乘着救护车巡回医疗的、做医生的父亲四处奔波.后来他和母亲、妹妹一起,在极端困难的情况下,到阿兰西去看他的外祖父母出了什么事.他曾在漫长的学校假期中,在外祖父宽敞的乡村花园里,度过了童年时代最幸福的时光.阿兰西靠近圣普里瓦的战场,为了到达这个城镇,他们三人不得不“在冰冻的严寒中”穿过被烧毁了的荒芜的村庄.最后他们到了目的地,只是发现家里已经被抢劫一空,“不只是有价值的东西,就连不值钱的东西”也被抢走了,并且被野兽般的行径弄得很脏,这种行径法国人在1870年和后来的1914年是很熟悉的.外祖父母什么也没有了;他们目睹大洗劫的那天,他们的晚饭是一位贫穷的妇女供给的,她不肯舍弃她已成废墟的农舍,并且坚持要与他们分食那一点点晚饭.

庞加莱永远忘不了这件事,他也忘不了南锡的长期被敌人占领.正是在战争期间,他学会了德语.庞加莱无法得到任何关于法国的消息,又急于想知道德国人关于法国和他们自己说了些什么,于是他自学了这种语言.他看见的和他





从入侵者本身的官方报道中知道的一切,使他终生成为一名热情的爱国者.但是,正像埃尔米特一样,他从来没有把他国家的敌人的数学,和他们更实际的行动混淆起来.另一方面,他的堂兄弟雷蒙只要一说到德国人的任何东西就恨得大叫大嚷.在那平衡一个爱国者的仇恨与另一个爱国者的仇恨的簿记上,庞加莱可以与库默尔相抵,埃尔米特与高斯相抵,这样就产生了隐含在圣经契约"以眼还眼,以牙还牙"之中的完美的零.

按照法国通常的习惯,庞加莱在从事专业之前参加了取得他的第一个学位(文学士及理学士)的考试.他在 1871 年 17 岁通过了这些考试——数学差一点没有通过!他迟到了,考试时有些慌,在收敛几何级数求和公式的极其简单的证明上失败了.但是他的名气已经走在他前面,主考人宣称:"除庞加莱以外,其他学生本来就应该得不及格的."

接着他准备了林学院的入学考试,在这所学校中,他使他的同伴们大吃一惊:他赢得了数学头奖而从不费事去做任何课堂笔记.他的同班同学们预先考过他,他们认为他是一个吊儿郎当的人,选出一个四年级的学生,用一道看起来特别难的数学难题来测试他.庞加莱不假思索地立刻给出解答就走开了,留下想引他上当的人问道:"他是怎么解出来的呢?"在庞加莱的一生中,其他一些人也会问同样的问题.当他的同事们交给他一道数学难题时,他好像从不思考,"答案像一支箭似地飞了出来".

这一年年底,他考入综合工科学校,名列第一.有几个关于他独特的考试的传说流传了下来.其中的一个讲到有一个主考人,事先得到庞加莱是一个数学天才的警告,就把考试暂停了 45 分钟,以便设计"一道'好'题"——一道精心想出的难题.但是庞加莱战胜了他,这位提问人"热烈地祝贺考生,告诉他,他赢得了最高分".庞加莱与折磨他的人一起的经历似乎表明,法国的数学主考人自从毁了伽罗瓦,又差一点同样毁了埃尔米特以来,已经学到了一些东西.

在综合工科学校,庞加莱以他在数学上的卓越才气,在所有体育锻炼,包括体操和军事训练上的极端无能,以及他完全没有绘画能力,画出的东西什么也不像而闻名.最后一点决不是开玩笑:他在入学考试时绘画得了零分,差点把他关在校门外面.这曾使他的主考人大伤脑筋:"……有了零分是不能录取的.在其他一切科目(除了绘画)上他都是绝对无可匹敌的.如果他被录取,那将是第一名;但是他能被录取吗?"既然庞加莱是被录取了,好心的主考人们可能在零前面放了一个小数点,在零后面添了个 1.

尽管庞加莱在体育锻炼上很无能,但在班上却很得人心.有一年年底,他们举办了庞加莱艺术作品的公开展览,用希腊文仔细地给展品标上标题,"这是一匹马",等等——并不总是准确的.但是庞加莱对绘画的无能,在他学习几何时表现出了严重的一面,他失去了第一的位置,在学校名列第二了.

庞加莱在1875年21岁时离开了综合工科学校，进了高等矿业学校，打算当一名工程师.虽然他诚心诚意地学习他的专业，但还是留下了一些空暇去研究数学，他在着手解决微分方程的一个一般问题上，显示了他的能力.3年后他向巴黎的科学院提交了一篇同一题目的论文，作为数学博士论文，但是这篇论文涉及一个更困难也更一般的问题.被请去审查这项工作的达布说："第一眼我就清楚地看出这篇论文是不同寻常的，很有接受的价值.它所包含的结果肯定足以为几篇好论文提供材料.但是，我必须不怕说出，如果要求庞加莱对他的工作方法提出一个精确的想法，那么很多地方需要修改或解释.庞加莱是一个直观主义者.一旦达到了顶峰，他从不追溯他的步骤.他满足于闯过难关，把勘测注定更容易通向尽头的坦途①的辛苦留给别人.他很乐意地做了我认为是必要的修正和整理.但是当我要求他做这件事时，他向我解释说，他脑子里有许多其他想法；他已经在忙着一些大问题了.他会把解答给我们的."

因此年轻的庞加莱就像高斯一样，被围攻他的头脑的大群想法压倒了，但是与高斯不同的是，他的座右铭不是"少些，但是要成熟".一个有创造力的科学家把他的劳动果实储藏了很久，以至它们中的一些已不新鲜了；比较性急的人则把采集到的一切不论生熟都散播出去，随着风和气候带它们落到可能成熟、可能腐烂的地方去；前一个人对科学的进展所做的工作是否比性急的人更多呢，这是一个没有解决的问题.有人认为这样，有人认为那样.由于做出一个决定超出了客观标准的范畴，人人都有权提出他自己的纯主观的意见.

庞加莱注定不会成为一个矿业工程师，但是他在见习期间表明了他至少有真正工程师的勇气.在一次造成16人牺牲的矿井爆炸和着火之后，他立即跟着救援人员下井了.但是这个职业同他的志趣不合，他喜欢他的学位论文和另一项早期工作为他开启的成为专业数学家的机会.他的第一次学术任命是1879年12月1日在卡昂(Caen)担任数学分析教授.两年以后(他27岁时)，他被提升到巴黎大学，1886年他再次晋升，在巴黎大学负责力学和实验物理课程(后一项看来有点奇怪，是考虑到庞加莱做学生时在实验室中的功绩吧).庞加莱除了去欧洲参加科学会议和1904年作为圣路易博览会邀请的演讲人去美国访问过一次以外，一直住在巴黎，成为法国数学的统治者.

庞加莱的创作时期开始于1878年的学位论文，终止于他1912年逝世——当时他正处在他力量的顶峰.他给这个相对短促的34年跨度塞进了大量著作，当我们考虑到大部分工作的难度时，其数量简直是绝对难以置信的.他的记录是近500篇关于新数学的文章，其中很多是范围广泛的研究报告，还有30多部实际上包括了数理物理学、理论物理学和理论天文学当时的所有分支的著作.

① 据说当亚历山大大帝想要很快地征服几何时，米内克穆斯(Menaechmus)对他说："几何无坦途."



这还没有计入他关于科学哲学的那些名著和他的通俗文章.谁要想对这个庞大的劳动量有一个恰当的概念,就必须成为第二个庞加莱,所以我们现在只从他的最驰名的著作中选出两三种,加以简单论述,谨在此最后一次为必需的不全面而道歉.

庞加莱的第一次成功是在微分方程理论方面,他把分析学的全部方法应用于微分方程,他绝对是分析学的大师.选择这个早期的主要努力方向,已经表明庞加莱对于数学的应用倾向,因为自从牛顿时代以来,微分方程吸引了众多的工作者,主要是因为它们在物理世界的探索中具有的重要性."纯"数学家有时候喜欢想象他们的一切活动都服从于自己的趣味,科学上的应用没有向他们提供什么有趣的东西.然而纯数学家中的一些最纯粹的数学家,终生孜孜不倦地致力于微分方程,这些方程首先出现在将物理状况转变成数学符号之中,实际上恰恰是这些人提出了这一理论的核心微分方程.由科学提出的一个特殊的方程,可以被数学家们推广,然后转回到科学家手里(经常是没有他们能够使用的任何形式的解答),以便应用到新的物理问题上,但是总的说来,动机是科学的.傅里叶的一段名言总结了这个论点,它使一类数学家恼怒,但是庞加莱赞成这种说法,并在他的大部分工作中遵循它:

"对自然的深入研究,"傅里叶宣称,"是数学发现的最丰厚的源泉.这种研究通过提出一个供研究的明确目标,不仅有排出含糊的问题和无用的计算的优点,而且也是一个形成分析本身和发现分析中那些原理的可靠方法,了解这些原理是必要的,科学总应该保存它们.这些基本的原理,就是在自然现象中反复出现的那些原理."对此,一些人可能会反驳:这很可能,但是在高斯意义上的算术怎么办呢? 无论如何,庞加莱是遵循了傅里叶的忠告,不管他是否相信它——甚至他在数论中的研究,也多多少少受到其他一些更接近于物理科学的数学研究的间接鼓舞.

庞加莱在微分方程上的研究,在1880年,他26岁时,导致了他的第一个最好的发现,即椭圆函数(和一些其他函数)的推广.三角函数 $\sin z$ 有周期 2π,即 $\sin(z+2\pi)=\sin z$;那就是说,当变量 z 增加 2π 时,z 的正弦函数回到它的初值.一个椭圆函数,比如说 $E(z)$,有两个不同的周期,比如说 p_1 和 p_2,使得 $E(z+p_1)=E(z)$,$E(z+p_2)=E(z)$.庞加莱发现周期性只是某种更普遍的性质的特例:当变量由它自身的可数无限多个线性分式变换之一代替时,某些特定函数的值还原,所有这些变换形成一个群.几个符号就能讲清楚这个论断.

设 z 被 $(az+b)/(cz+d)$ 代替,那么对于 a,b,c,d 之值的某个可数无限集,有 z 的一些单值函数,比如说其中之一为 $F(z)$,使得

$$F\left(\frac{az+b}{cz+d}\right)=F(z)$$

进而，如果 a_1, b_1, c_1, d_1 和 a_2, b_2, c_2, d_2 是 a, b, c, d 的值集中的任意两个，又如果 z 先被$\frac{a_1 z + b_1}{c_1 z + d_1}$代替，然后在这个式子中，$z$ 被$\frac{a_2 z + b_2}{c_2 z + d_2}$代替，得到比如说$\frac{Az + B}{Cz + D}$，那么我们不仅有

$$F\left(\frac{a_1 z + b_1}{c_1 z + d_1}\right) = F(z), F\left(\frac{a_2 z + b_2}{c_2 z + d_2}\right) = F(z)$$

而且有
$$F\left(\frac{Az + B}{Cz + d}\right) = F(z)$$

更进一步，如刚才解释的那样保持 $F(z)$的值不变的所有置换

$$z \to \frac{az + b}{cz + d}$$

(箭头读作"用……代替")的集合形成一个群：集合中两个置换相继实施的结果

$$z \to \frac{a_1 z + b_1}{c_1 z + d_1}, z \to \frac{a_2 z + b_2}{c_2 z + d_2}$$

仍在集合中；集合中有一个"恒等置换"，即 $z \to z$(这里 $a = 1, b = 0, c = 0, d = 1$)；最后，每一个置换有一个唯一的"逆"——就是说，对于集合中的每一个置换，有一个单独的另一个置换，如果把它作用到第一个置换上去，就产生恒等置换．总之，用前面的术语，我们看出 $F(z)$是一个在一个线性分式变换的无限群下不变的函数．注意，如开始说的那样，置换的无限是可数的无限：置换能用 1，2，3，…数出来，不像直线上的点那样多．庞加莱实际构造出了这样的函数，他在 19 世纪 80 年代的一系列文章中发展了它们的最重要的性质．这样的函数称为自守函数．

这里只需说两点，以指出庞加莱用这个奇妙的创造取得了什么样的成就．首先，他的理论把椭圆函数理论作为一个细节包括进去．其次，正如著名的法国数学家乔治·安贝尔(Georges Humbert)所说，庞加莱发现了两个值得注意的命题，这些命题"给了他代数和谐的钥匙"：

在同一个群下不变的两个自守函数①，是由一个代数方程联系起来的；

反之，在任何代数曲线上的某个点的坐标，都能够用自守函数来表示，因此可用一个参数(变量)的单值函数来表示．

一条代数曲线是其方程具有类型 $P(x, y) = 0$ 的曲线，其中 $P(x, y)$是 x, y 的多项式．举一个简单的例子，中心在原点——(0,0)——半径为 a 的圆的方程是 $x^2 + y^2 = a^2$．按照庞加莱的第二把"钥匙"，把 x, y 表示成一个单参数的，比如说 t 的自守函数，一定是可行的．它确实是可行的，因为如果 $x = a\cos t$，

---

① 庞加莱以现代微分方程理论的创始者之一德国数学家拉扎勒斯·富克斯(Lazarus Fuchs, 1833—1902)的名字，称他的一些函数为富克斯函数，其原因无须在这里说明．他以费利克斯·克莱因的名字称其他一些函数为克莱因函数——讽刺性地承认有争议的优先权．





$y = a\sin t$,那么,开平方并相加,我们就消去了 t(因为 $\cos^2 t + \sin^2 t = 1$),得出 $x^2 + y^2 = a^2$.但是三角函数 $\cos t$,$\sin t$ 是椭圆函数的特例,而椭圆函数又是自守函数的特例.

这个广阔的自守函数理论的创立,只不过是庞加莱在30岁以前做出的分析学中许多令人吃惊的东西之一.他并没有把全部时间都奉献给分析学,数论、代数的一些部分,数理天文学,占据了他的注意力.在数论中,他以几何形式重新建立了双二次形式的高斯理论,这特别合乎那些像庞加莱一样喜欢直观方法的人的心意.这当然不是他在高等算术理论中做过的全部贡献,但是限于篇幅,不能进一步详述.

这个本事很大的人的工作,从没有不受人重视.在32岁(1887年)这个不寻常的年轻年龄,庞加莱被选入科学院.他的提名人说了一些言过其实的话,但是大多数数学家会同意这是真实的:"(庞加莱的工作)高于通常的赞扬,而且使我们不可避免地想起雅可比写的有关阿贝尔的话——他解决了在他之前无法想象的问题.确实必须承认,我们正目睹一场数学上的革命,这场革命在每个方面都可以与半个世纪以前,由于椭圆函数的出现所产生的革命相比."

在这里离开庞加莱的纯数学工作,就像刚刚坐在宴会桌旁就站起来一样,但是我们必须转到他的多方面才能的另一方面了.

自从牛顿和他的直接后继者的时代以来,天文学慷慨地给数学家们提供了比他们能解决的还要多的问题.直到19世纪后期为止,数学家们在着手解决天文学问题时所用的方法,实际上都是由牛顿、欧拉、拉格朗日和拉普拉斯发明的那些方法的直接改进.但是贯穿整个19世纪,特别是自从柯西对单复变量函数的发展,以及他本人和其他一些人关于无穷级数收敛性的研究以来,纯数学家的工作中已经积聚了大量未经试验的方法.对于庞加莱,分析对他来说像思考一样自然,这一大堆未曾用过的数学,好像是世界上用来解决天体力学和行星演化的突出问题的最自然的东西.他从这一大堆中挑拣了他喜爱的东西,改进了它们,发明了他自己的新方法,以一个世纪以来理论天文学没有经受过的庞大方式向它进攻.他把这个进攻现代化了;对于大多数天体力学专家来说,他的战役确实是极端现代的,以至于甚至今天,在庞加莱开始他的进攻40年或更久以后,还很少有人掌握他的方法,一些不能运用他的方法的人,暗示它在实际解决问题上是没有价值的.然而庞加莱并不缺乏有力的拥护者,他们的胜利对于庞加莱时代以前的人们是不可能的.

庞加莱在数学天文学上的第一次(1889年)成功,出自他对"n体问题"的一个不成功的考虑.对于 $n = 2$,这个问题已由牛顿完全解决了;著名的"三体问题"($n = 3$)将在后面谈到;当 n 超过3时,一些可以适用于 $n = 3$ 的情形的化简可以继续下去.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

根据牛顿的万有引力定律，两个质量为 m, M，间隔距离为 D 的质点，以与 $(m \times M)/D^2$ 成正比的力相互吸引．想象空间中随意分布的 n 个质点；假定所有质点的质量、初始运动和相互之间的距离在某一给定的瞬间都是已知的．如果它们按照牛顿定律互相吸引，那么在任意规定的时间间隔之后，它们的位置和运动(速度)是什么呢？对于数理天文学，在一个星团、或一个星系、或一个星系团中的恒星，可以看做是按牛顿定律互相吸引的质点．这样，“n 体问题”就相当于——在它的一个应用中——问从现在起 1 年或 10 亿年后，天空是什么样子的，人们认为，我们有足够的观测数据，可以描述今天的总体位形．当然问题由于辐射而极其复杂——恒星的质量不会历经数百万年而不变；但是在牛顿形式下的 n 体问题的某个完全的可计算解，也许会给出对人类的一切目的而言足够精确的结果——人类很可能在辐射能够造成可以观测到的误差之前很久就消亡了．

这实质上是为瑞典国王奥斯卡二世(King OscarⅡ)在 1881 年提供的奖金所提出的问题．庞加莱没有解决这个问题，但在 1889 年，由于他对动力学微分方程的一般讨论和对三体问题的研究，由魏尔斯特拉斯、埃尔米特和米塔－列夫勒组成的评奖团授予了他这一奖金．三体问题通常被认为是 n 体问题最重要的情形，因为地球、月球和太阳提供了一个 $n=3$ 的情形的例子．魏尔斯特拉斯在给米塔－列夫勒的报告中写道：“您可以告诉您的国王，这项工作确实不能看做提供了所提问题的完全解答，但是不管怎么说，它是非常重要的，它的发表将在天体力学史上开创一个新的时代．因此可以认为陛下在设立这一项竞赛时所考虑的目的已经达到．”法国政府不甘心落在瑞典国王后面，在庞加莱获得奖金后封他为荣誉军团骑士——这是对这位年轻天才数学家所给的比国王的 2 500 克朗和金质奖章便宜得多的奖励．

由于已经提到过三体问题，现在可以报告它最近的一项发展；自从欧拉那时以来，它就被认为是整个数学领域中最困难的问题之一．从数学上讲，这个问题归结为解九个微分方程(都是二阶线性的)的联立方程组．拉格朗日成功地把这个方程组约化为更简单的形式．如同在大部分物理问题中那样，无法期望有限形式的解；如果确实存在一个解，它将由无穷级数给出．如果这些级数(形式上)满足那些方程，而且对于变量的一些值收敛，那么解就“存在”，中心困难是证明收敛．到 1905 年为止，已经发现了各种各样的特殊解，但是尚未证明存在任何能被称为通解的东西．

在 1906 年和 1909 年，一个完全意想不到的地区——芬兰——取得了一项重要进展，对于这个国家，老于世故的欧洲人甚至今天还认为它几乎是不开化的，特别是认为它还债的怪习不够文明；在帕沃·努尔米①在径赛中跑过美国之

① 帕沃·努尔米(Paavo Nunni，1897—1973)为芬兰径赛运动员．在 3 次奥运会上共获 6 枚个人金牌，3 枚团体金牌，是奥运会史上第一个 9 枚金牌获得者．——译者注





前,很少有美国人认为这个国家超出了石器时代.除了所有三体同时撞到一起的极罕见的情形,赫尔辛福斯①的卡尔·弗里肖夫·宗德曼(Karl Frithiof Stmdman)利用归功于意大利人莱维·齐维塔(Levi Civita)和法国人潘勒韦(Painlevé)的分析方法,做出他自己的巧妙变换,证明了存在上述意义的解.宗德曼的解答不适宜于数值计算,也没有提供多少关于实际运动的信息,但那不是这里感兴趣的事:一个不知道是否可解的问题被证明是可解的了.许多人曾经不顾一切地努力证明这个重要结论,当证明即将出现时,一些数学家像常人常有的那样,急忙指出宗德曼没有做出什么重要的事,因为除了他解决的那个问题以外,他没有解决其他的问题.这一类批评在数学上就像在文学和艺术上一样平常,这再次表明数学家同任何人一样是凡人.

庞加莱在数学天文学中最有独创性的工作,总结在他的伟大专著《天体力学新方法》(三卷,1892、1893、1899)中.接着是1905~1910年的具有更直接实用性质的另一部三卷集著作《天体力学教程》,以及再晚一些出版的他的课程讲义《流体质量平衡的图形》和一本历史性评论著作《关于宇宙论假设》.

达布(得到其他许多人的支持)就这些著作中的第一部发表看法,说它确实开创了天体力学的一个新时代,它可以与拉普拉斯的《天体力学》和达朗贝尔早期关于岁差的著作媲美."沿着拉格朗日开创的分析力学的道路,"达布说,"……雅可比建立了一个看来是动力学中最完整的理论.50年来,我们依靠这位杰出的德国数学家的那些定理,从各种角度应用它们,研究它们,但是没有添加任何本质的东西.正是庞加莱第一次打破了似乎装着这个理论的这些僵硬的框子,为它设计出客观世界的前景和新的窗口.他在动力学问题的研究中引进或使用了不同的概念:首先是变分方程,即决定某一问题的无限接近一个已知解的解答的线性微分方程,这个概念以前就有过,而且不仅仅可以应用于力学;其次是积分不变式,这完全属于他,并在这些研究中起了重要作用.再加上其他一些基本概念,特别是涉及所谓"周期"解的那些概念,对于它们,被研究其运动的那些物体,在一定时间后回到它们的初始位置和初始相对速度."

周期解的概念开创了数学的整整一个部门,即对周期轨道的研究:比如说,给定一个行星系统,或一个恒星系统,以及该系统中所有成员在某一确定时刻的初始位置和相对速度的完整数据,要求确定在什么条件下该系统会在稍后的某一时刻回到它的初始状态,从此无限地重复它的循环运动.例如,太阳系具有这种周期类型吗?或者如果不具有,那么要是它是孤立的,而且不受外界天体的摄动,它会具有吗?无须说,这个一般的问题还没有完全解决.

庞加莱的许多天文学研究工作是定性的,而不是定量的,正如一个直觉主

① 即赫尔辛基.——译者注

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

义者应有的那样,这个特点把他引向了拓扑学研究,就像它曾引领过黎曼一样.他发表了六篇关于拓扑学的著名论文,它们革新了当时的这个学科.关于拓扑学的研究又被大量应用于天文学的数学.

我们已经提到庞加莱关于旋转流动物体问题的工作——在天体演化学中具有明显的重要性,它的一个方面假定行星一度很像这样一些流动物体,它们可以被看做好像它们实际上不是那么荒谬可笑的样子.它们是否如此对这种情形的数学而言并不重要,它本身就很有意思.从庞加莱自己的概述中摘出的几段话,比任何释义更清楚地说明了他在这个困难的学科中用数学说明的东西的性质.

“让我们想象一个(旋转)流动物体由于冷却而收缩,但是收缩慢得足以保持均匀,并使旋转在一切部分都一样.

“一开始,形状很近似一个球的这团物体变成一个旋转椭球,它将变得越来越扁,然后,在某个确定的时刻,它将变成一个有三个不同轴的椭球.再后来,形状不再是椭球,而成了梨形,直到最后这一团物体在它的‘腰’部越来越凹进去,分成两个隔开的、不同的物体为止.

“上述假说肯定不能应用到太阳系.一些天文学家认为它可能对一些双星是成立的,天琴座β型双星可能会呈现出类似于我们讲到的那些形式的过渡形式.”

然后他继续提出该研究对土星光环的应用,他宣称已经证明了只要光环的密度超过土星密度的1/16,光环就是稳定的.可以说,直到1935年这些问题还不能被看做完全解决了.特别是对可怜的老土星的更严格的数学处理,似乎表明它还没有被大数学家们完全征服.这些数学家中包括克拉克·麦克斯韦,过去的70年中这些数学家一直在断断续续地不停地研究它.

我们必须再次离开这个几乎还没有尝到什么东西的宴会,进入庞加莱在数理物理学中的大部头著作.这里他的运气不是那么好.要想乘机利用他自己卓越的才能,他应该晚生30年或多活20年.不幸的是,当他正处在全盛时期时,物理学正好经过一个循环之后进入衰老期;当物理学开始恢复青春时——在普朗克于1900年,爱因斯坦于1905年,完成了赋予这个衰老的浪子第一对新命脉的困难棘手的手术之后——他已经如此彻底地沉浸在19世纪的理论中,以至于他在1912年去世之前几乎没有时间去领悟这些奇迹;庞加莱在他整个成年时期,似乎都是通过思考去汲取知识,而没有作有意识的努力.他像凯莱一样,不只是一个多产的创造者,而且也是一个学识渊博的学者.他的范围也许比凯莱还要宽广,因为凯莱从来没有自称他能够懂得在应用数学中正在发生的一切.这种独特的博学,当碰到与经典科学相对立的、现实的科学问题时,也许是一个不利因素.





一切在物理学坩埚中沸腾的东西只要一出现,庞加莱就会立刻掌握它们,把它们构成几个纯数学研究的题目.当无线电报发明时,他抓住这个新东西,创立了它的数学.当其他人或者无视爱因斯坦关于(狭义)相对论的早期工作,或者把它只作为一个奇怪的东西忽略过去时,庞加莱已经在忙着它的数学了,他是第一个告诉世界什么已经到来的赫赫有名的科学界人士,他催促全世界把爱因斯坦的理论看做也许是新纪元的最重要的现象,他预见到这个新纪元,但是他自己不能开创这个新纪元.对于普朗克量子理论的早期形式也是如此.当然,有各种不同的意见;但是时间过得这样久了,现在看来,数理物理学之于庞加莱,犹如谷神星之于高斯;虽然庞加莱在数理物理学方面所完成的工作足够赢来半打伟大的荣誉,但那不是他毕生从事的职业,要是他坚持从事纯数学,科学从他那得到的就会更多——他的天文学工作也就是纯数学,而不是其他.但是科学得到的已经够多了,而且一个具有庞加莱的天才的人,有权做他喜爱的事.

我们现在还有篇幅来讲述庞加莱多才多艺的最后一个方面:他对数学创造的基本原理的兴趣.在 1902 年和 1904 年,瑞士数学杂志《数学教学》对数学家们的工作习惯进行了调查.调查表发给了许多数学家,其中一百多人作了回答.对于问题的回答和对一般倾向的分析,在 1912 年以最后形式发表①.任何希望一窥数学家的“心理”的人,都能在这件别具一格的著作中,以及庞加莱在看到这个调查的结果之前独立得出的观点的证据中,找到许多有趣的东西.在我们引用庞加莱的话以前,可以提到普遍感兴趣的几点.

那些后来成为大数学家的人对数学的早期兴趣,已经在前面多次举例说明了.对于“在什么时期……和在什么情况下数学抓住了你?”这个问题,收到了 93 人对第一个提问的回答:35 人说在 10 岁以前,43 人说在 11 岁到 15 岁;11 人说在 16 岁到 18 岁;3 人说在 19 岁到 20 岁;只有 1 个晚的说在 26 岁.

另外,凡是有数学家朋友的人都会注意到,他们当中有一些人喜欢在早晨很早的时候工作(我知道一个很出众的数学家,他在早得不近人情的清晨 5 点钟开始他一天的工作),而另外一些人在天黑以前什么都不做.对于这一点的回答,表明了一个奇怪的倾向——也许是重要的,尽管有许多例外:北方民族的数学家们宁愿在夜里工作,而拉丁系民族的数学家喜欢在早晨工作.在夜里工作的数学家中,长时间的集中精力,使得他们年纪大一些的时候常常失眠,于是他们——不情愿地——改成早晨工作.费利克斯·克莱因在年轻时夜以继日地工作,他有一次指出摆脱这种困境的一个可能的途径.他的一个美国学生抱怨由于思考数学而不能睡觉.“睡不着,嗯?”克莱因哼了一声,“水合氯醛是干什么的?”不过,不能不加区别地推荐这种药物;克莱因自己悲剧性的衰竭也许与它

---

① 《数学教学》关于数学家工作方法的调查,可从该杂志及巴黎的戈蒂埃-维拉尔出版社的单行本中见到.

有关.

也许,收到的回答中最重要的是,那些关于灵感与苦干哪一个是数学发现的源泉这一问题的回答.结论是"数学发现,不论大小……都永远不会自发地产生.它们总是自觉和不自觉地以一片播种了基础知识,并通过劳动充分耕耘的土壤为先决条件的."

那些宣称天才基于百分之九十九的汗水,只有百分之一的灵感的人,如托马斯·阿尔瓦·爱迪生(Thomas Alva Edison)和那些把这个数字颠倒过来的人并不抵触.哪方都是对的.一方记住了苦干,而另一方在貌似突然发现的激动中把它忘得精光,但是当他们分析他们的印象时,双方都承认,没有苦干和"灵感"的闪现,就不可能有发现.如果只是苦干就够了,那么许多酷爱艰苦工作的人,他们对于某个科学分支似乎无所不知,同时又是很好的批评家和评论家,怎么会连一个小小的发现也从来没有做出来呢?另一方面,那些相信"灵感"是科学或文学上发现和发明的唯一因素的人,在看了雪莱(Shelley)的任何一首"完全自发的"诗早先的草稿(只要这些草稿保存了下来,并被复制了),或者巴尔扎克使他的气得发狂的出版商承受的任何一部长篇小说的一次次修改稿后,就可能发现这是很有教益的.

庞加莱在一篇 1908 年发表,以后又收入他的《科学方法》中的文章里,说明了他对数学发现的看法.他说,数学发现的产生,是一个应该使心理学家们非常感兴趣的问题,因为它是人类头脑似乎从外部世界借用得最少的活动,而且通过了解数学思维的过程,我们有希望得知什么是人类头脑中最本质的东西.

庞加莱问道:"怎么竟会有一些不懂数学的人呢?这应该使我们感到吃惊,或者,更确切地说,如果我们不是如此的对它习以为常,那么它就会使我们吃惊."如果数学仅仅是建立在逻辑规则的基础上,如同所有正常的人都接受的那样,而且(按照庞加莱的说法),只有疯子才会否认这一点,那么,怎么会有这么多人不懂数学呢?对这个问题也许可以这样回答,还没有哪一组证实数学上的无能是正常人类模式的详尽试验公布过."还有,"他问:"在数学中错误怎么会是可能的呢?"亚历山大·蒲柏(Alexander Pope)说:"犯错误是合乎人性的."这是和任何其他解答一样不能令人满意的解答.消化系统的化学也许与它有点关系,但是庞加莱想要更微妙的解释——一个不能以给"无价值的躯体"喂麻醉剂和酒精来做试验的解释.

"在我看来答案是明显的."他宣称.逻辑同发现或发明没有什么关系,记忆力起了作用.不过记忆力并不像它可能的那样重要.他一点不脸红地说,他的记忆力不好:"那么它为什么没有在一段困难的、大多数棋手(他认为他们的"记忆力"是极好的)都会失败的数学推理中抛下我呢?显然是因为它受到推理的一般进程的指导.一个数学证明不只是演绎推理的并列;它是按一定顺序排列的





演绎推理,而顺序是比组成部分本身更重要的."如果他有这个顺序的"直觉",记忆力就算不得什么了,因为每一个演绎推理都将自动占据它在序列中的位置.

然而数学创造不仅在于做出已知事物的新组合;"任何人都能做出组合,但是这样的组合会有无限多个,大多数是没有意义的.创造恰恰在于避免无用的组合,做出那些有用的组合,它们只构成一小部分.发明是识别、选择."但是这一切不是以前就说过几千遍了吗?有哪一个艺术家不知道选择——一个不可捉摸的东西——是成功的一个秘密呢?我们恰恰处于我们在开始研究以前的地方.

为了结束庞加莱的这一部分观察报告,可以指出,他说的许多东西都是建立在一个假定上的,这个假定可以确定是成立的,但是它没有一点儿科学证据.直截了当地说,他假定人类的大多数都是数学上低能的人.如果在这点上同意他,那么我们甚至无须接受他的纯浪漫的理论.它们属于灵感的文学,而不属于科学.我们转到一些不那么引起争论的东西,现在来引用一段著名的文字,庞加莱在这段文字中描绘了他自己的一个最伟大的"灵感"是怎样产生的,用意在于证明他的数学创造的理论.它是否起到了这个作用,可以留待读者去判断.

他首先指出,为了领会他的叙述,不需要懂得专业术语:"心理学家们感兴趣的不是定理,而是环境."

"我花了 15 天,竭力证明不存在与我后来称为富克斯函数类似的函数,我那时非常无知.每天坐在我的工作台前,花上一两个小时的时间;我试了很多种组合,什么结果也没有得到.一天傍晚,与我的习惯相反,我喝了黑咖啡;我不能入睡,各种想法蜂拥而至;我觉察到它们互相冲突,直到一对想法,就这么说吧,钩在一起,形成一个稳定的组合.到早晨我已经证实了一类富克斯函数,即从超几何级数中得出的那类函数的存在.我只需要写出结果,这用了我几个小时.

"接着我想用两个级数的比来表示这些函数;这个想法是完全有意识的,想出来的;与椭圆函数的类比指引了我.我问自己,如果这些级数存在的话,它们的性质是什么,我毫无困难地构造出了这些级数,我称它们为富克斯级数.

"然后我离开了当时我住的地方卡昂,参加矿业学院主办的一次地质调查旅行.旅行的紧张使我忘记了我的数学劳动;到达库唐斯以后,我们乘公共汽车去某地游览,就在我踏上汽车踏板的那一瞬间,产生了一个想法,显然我以前并没有想到什么东西为它作准备,这个想法是,我用来定义富克斯函数的那些变换,与非欧几何的那些变换是一致的.我没有作证明;我不会有时间,因为一上车,我就在继续一个中断了的谈话;但是我立即感到这是完全确定无疑的.一回到卡昂,就在我的空闲时间证明了这个结果,以使自己安心.

"然后我着手研究一些算术问题,没有什么明显的成绩,也没有怀疑过这些

问题会与我以前的那些研究有任何联系.研究得不到成功,使我感到厌烦了,于是我去海边待了几天,思考另一个问题.一天,当我沿着海边的峭壁散步时,那个想法又出现了,又带有那种简明、突然、立即确定的特点,这就是三元二次不定型的那些变换,与非欧几何的那些变换是一致的.

"一回到卡昂,我就考虑这个结果,得出了它的结论;二次型的例子向我表明,有一些富克斯群与相应于超几何级数的那些富克斯群不同;我看出我能把它们应用于 θ 富克斯函数的理论,因此存在一些与从超几何级数导出的那些 θ 富克斯函数不同的 θ 富克斯函数,这是我到那时为止所知道的唯一不同的 θ 富克斯函数.很自然地,我给自己提出了构造所有这些函数的任务.我系统地做出了再三努力,一个接着一个地解决了所有的外围问题;然而有一个问题仍然解决不了,而攻下它就会取得全盘胜利.但是我的全部努力只是使我更通晓了这个问题的困难,而这本身是有意义的.所有这些工作都是完全有意识地去做的.

"正在这时,我离开卡昂前往瓦勒里昂山,在那里服兵役.因此我有了一些很不同的急务.一天,在穿过大街时,挡住我的那道难题的解答突然出现了.我没有试图立刻深入研究它,只是到了我的服役结束以后,我才继续研究这个问题.我已经有了所有的组成部分,只需要装配它们,排列它们就行了.所以我写出了我最后的论文,一挥而就,毫无困难."

这类事情,他说,还可以从他自己的工作和其他数学家的工作中,举出许多其他的例子,如在《数学教学》杂志中报告的.从他的经历中,他相信这种表面上的"突然启发","(是)以前长期潜意识工作的明显标志",他继续推敲他的潜意识思维理论和它在数学创造上的作用.有意识的工作,作为引发潜意识长期积聚的炸药的一种触机是必要的——他没有这样说,但是他所说的与此相当.但是如果我们仿效庞加莱,把正是我们要了解的那些活动,强加在"潜意识思维"或"潜在自我"之上,那么我们能从理性解释的方式中得到些什么呢?不是赋予这个神秘的作用者一种假设的辨别力,使它能够在提出来(怎样提出来,庞加莱没有说)供它检查的"数目极多的"可能组合中做出区别,而是平静地说"潜意识"除了"有用的"组合外,拒绝一切组合,因为它对对称和美有一种感受力,这听起来不免像是给人一个印象较深的名字来解决那个初始问题.也许这正是庞加莱打算做的,因为他一度把数学定义为给各种不同的事物同一名字的艺术;所以在这里他可能是通过给同一事物以不同的名字来完成其观点的对称性.看来很奇怪,一个能够满足于数学创造的这样一种"心理学"的人,在宗教问题上竟然是一个完全的怀疑主义者,庞加莱就是这样一个人.在庞加莱出色地陷入心理学之后,怀疑主义者们很可能会不再对不相信任何东西抱什么希望了.

在 20 世纪头 10 年间,庞加莱的名声增长得很快,他开始——特别是在法





国——被看做一切关于数学的事情的大智者.他对各种从政治到伦理学的问题的见解,通常是直截了当的,并被大多数人当做最终结论来接受.正像一个伟人去世后几乎必然会发生的那样,庞加莱生前令人眼花缭乱的名声,在他去世后的10年中经历了一个不公平的黯然失色的时期.但是他对于什么东西会使后代感兴趣的直觉已经在为它自己辩护了.只举许多例子中的一个,庞加莱是全部数学都能用经典逻辑的基本符号重写这一理论的有力反对者;他认为是某种超出逻辑的东西,使数学成了现在的样子.虽然他并未像现在的直觉学派走得那样远,但是他也像那个学派一样,似乎相信至少一些数学概念先于逻辑概念,并且如果两者之间有因果关系,那么正是逻辑必定出自数学,而不会反过来.这是否就是最终信条还有待考虑,但是目前看来,好像庞加莱以听凭他自由使用的一切冷嘲去攻击的理论,不管具有什么样的优点,并不是最后的理论.

在庞加莱的最后4年中,除了令人苦恼的疾病之外,他繁忙的生活是平静而幸福的.各种荣誉从全世界所有的主要学术团体那里雨点般地向他飞来.在1906年,52岁时,他获得了法国科学家可能得到的最高的荣誉称号——法兰西科学院院长.所有这一切并没有使庞加莱妄自尊大,因为他是真正谦逊单纯的,不装腔作势.他当然知道,在他的壮年时期他没有一个接近他的对手,但是他也能毫不做作地说,与要知道的东西相比,他什么都不知道.他很幸福地结了婚,有一个儿子和三个女儿,他从他们那里得到了很多乐趣,特别是在他们幼小的时候.他的妻子是艾蒂安·若弗鲁瓦·圣伊莱尔(Étienne Geoffroy Saint-Hilaire)的曾孙女,圣伊莱尔作为那个爱争吵的比较解剖学者居维叶(Cuvier)的对手而为人们所记忆.庞加莱的爱好之一是交响乐.

在1908年于罗马举行的国际数学大会上,庞加莱因病没有能宣读他那激动人心的(也许是过早的)演讲——数理物理学的未来.他的病是前列腺增大,意大利的外科医生们给他做了手术,解除了症状,人们以为他得到了根治.回到巴黎以后,他像以前一样精力充沛地继续工作.但是在1911年,他开始有了可能不久于人世的预感,12月9日他写信给一个数学杂志的编辑,询问是否能接受——与通常的习惯相反——一篇尚未完成的论文,关于庞加莱认为最最重要的某个问题的论文:“……以我的年纪,我可能不能解决它了,所得到的结果,有可能把研究者们带到新的、意想不到的道路上去,尽管它们使我多次受骗,我认为它们太有前途了,我自愿献出它们……”他已经把两年中大部分时间用来试着去克服他的困难,但都徒劳无功.

他猜测的那个定理的证明,能够使他在三体问题上取得惊人的进展;特别是将使他能够证明比以前考虑过的更一般的某些情形的无限多个周期解的存在.这个期望中的证明,在庞加莱的“未完成交响曲”发表以后不久,就由一个年

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

轻的美国数学家乔治·戴维·伯克霍夫(George David Birkhoff,1884—)[1]证明了.

1912年春天,庞加莱再次病倒,7月9日接受了第二次手术.手术是成功的,但是7月17日,他在穿衣服的时候,因栓塞猝死.他当时59岁,正处在能力的顶峰——用潘勒韦的话说,是“理性科学活着的大脑”.

① 伯克霍夫于1944年逝世.——译者注





庞加莱和数学

第二章

1 庞加莱和数学①

1954年5月15日，在Sorbonne(索邦)隆重举行了庞加莱诞生一百周年纪念大会，共和国总统、教育部长、法国和各国著名的科学家参加了大会，Jacques Hadamard, Henri Villat, Louis de Broglie和Maurice de Broglie在会上发表了演说，分别介绍了庞加莱在数学、力学、物理学和哲学方面的工作，现将阿达玛的演说译载如下.

今天，法兰西在纪念她的民族骄子之一庞加莱.他的名字应该是人所共知的，应当像他生前在人类精神活动的另一个领域那样，使每一个法国人感到骄傲.数学家的业绩不是一眼就能看见的，它是大厦的基础、看不见的基础，而大厦是人人都可以欣赏的，然而它只有在坚实的基础上才能建立起来.

当一个人从事一件并不那么轻松的任务，如同今天我荣幸地承担的任务这样，要在几分钟之内，介绍一个伟大的事业、一个在人类智慧上划时代的事业时，理所当然地想从中发现共性，并从共性中找到个性，特别是，在一位独树一帜、才智超人的天才人物身上，他的个性就更加突出.

① 原载自：Oeuvres de Henri Poincaré Tome XI，郭书春，译，桂来强，校.

然而,我们今天没有必要这样做,如果这样做,我看就会贬低同时也会歪曲庞加莱的事业.

“我们宁作数学的仆人而不作数学的主人”,一天埃尔米特对我说.最常见的情况是,一位学者一般是根据自己的可能来选择他的课题的.一般水平的研究者可能如此行事.

这样做的是一般水平的研究者,而不是庞加莱.庞加莱不是根据自己的才能,而是根据科学的需要来确定其研究课题的.科学的需要就是他思想的出发点.他的思想可以说不是产生在他身上,而同时,是一种“超人的力量”,这是埃尔米特非常珍视的另一句话,在他身上产生了唯有他才能见到的光辉,而他使这光辉普照众人.

请允许我简单地回顾一下,他来到人间时,数学的状况是怎样的.

18 世纪留给 19 世纪两大难题:微分方程的积分法与偏微分方程的积分法.19 世纪的数学在这方面开辟了一条新的途径.当有规律地给变量不仅引入实数值而且也引入虚数值时,就把研究中出现的性质阐述得意想不到的清楚.

这种“函数论”,说得更确切些,这种“解析函数论”,在法国是柯西的功劳,而在德国则是魏尔斯特拉斯的功劳.他们分别创立了解析函数论,并且奠定了微分方程论的初步基础.

柯西(Cauchy, Augustin-Louis, 1789—1857),法国数学家、力学家.

柯西逝世于 1857 年.魏尔斯特拉斯在他之后还活了 40 年,在德国科学界是一位受到无限赞誉的人物.但是,他的最有名的弟子之一米塔 - 列夫勒告诉我们,在他的一生行将结束的时候,他并非毫无忧虑地看到,他曾使德国科学得到的领先地位,正在转到我国.

魏尔斯特拉斯(Weierstrass, kart Theodor Wilhelm, 1815—1897),德国数学家.

事实上,这是科学上刚刚发生的一件大事.自守函数从庞加莱的一系列论文中,用当时的说法,“爆发”出来了.

解析函数论最辉煌的成果之一曾经是椭圆函数论.由于庞加莱的大量推论、概括,同椭圆函数的性质一样美好的全部性质,被推广到无限多个实体,这些实体都属于同一个总的理论范畴,虽然它们彼此存在着深刻的差别.

米塔 - 列夫勒(Mittag-Leffler, Magnus Gustaf, 1846—1927),瑞典数学家.

人们曾经说过,自守函数是“打开代数世界的钥匙”,并且,在一种重要情况下,即代数系数线性方程的情况下,解决了我们开始时谈到的另一个重大问题,即微分方程的积分法.

还是在发表自守函数的论文之前,庞加莱就发表了几篇关于《算术》问题的简短注释.我要对其中在我看来能表明他某一方面特点的一篇说几句话.这篇注释专讲连续约简方法,即埃尔米特的著名发现,这种发现引起了我们的赞叹,然而,我们却不明白发现者是怎样想出这个方法的.

然而,在埃尔米特的著作中我们捉摸不透的东西,庞加莱把它阐述得非常清楚;并且,庞加莱的这种既深入浅出又直截了当的特点,始终如一地贯穿于他





的整个著作中.

那么,这是不是说,庞加莱同他所介绍的这位伟大的前辈相反,没有受到灵感的启示呢?不能这样说,因为在我们面前有关他的灵感的脍炙人口的故事,这些灵感给他生涯的开端打上了印记并且奠定了自守函数论的基础.此外,在《天体力学的新方法》的第三卷中,有一个说明无意识起作用的非常典型的例子,在那里,庞加莱不自觉地谈到了变分法.那时,魏尔斯特拉斯刚刚给予变分法一种严格的方法,这种方法给了问题以完美和最终的解答,只是由于魏尔斯特拉斯在他的晚年不屑于出版他的这些成果,以至于这些成果长期以来只有听过他讲演的人才知道.

庞加莱到底是已经知道,还是重新发现了魏尔斯特拉斯的这种方法呢?读过他的名著的人被弄糊涂了,在同一页中既有一个只能由对魏尔斯特拉斯的方法一无所知的人写出的句子,也有另一个只能由熟知魏尔斯特拉斯方法的人写出的句子.一个人俨然成了两个人,这种特征表明一次甚至两次干预都是无意识发生的,每个人都有其无意识于某种事情的方式,只是我们无法弄清其来龙去脉罢了.

然而,这种情况是绝无仅有的.尽管庞加莱在思维的时候,很可能是无意识的,但是他的思路却是十分清楚的,这给大家一个错觉,以为他可能就是这样做出发现的.

在自守函数论问世之后,数学家们对他的赞扬大概就没有断过.新的发现接踵而来,其速度之快,简直令人难以想象.这些发现构成 *Bulletin de La Société Mathématique de France*(法国数学会公报)的单独一册,即第 XI 卷,它包含三篇论文,这三篇论文在理论方面都大有创新.

在后来的岁月中,具有决定性意义的论文仍然以惊人的速度接连出现.

这些如此富有成果的研究包括哪些方面呢?

包括各个方面.也许不存在几何学家们不能解决的大问题,不存在天才的青年学者没有给我们指出道路的问题.

首先是解析函数论,直到那时,它的两个还未探索而又特别困难的重大课题.已经摆脱了一无所知的状态.对其他作者做出的头等重要的发现我们也是非常关心的,但是我们在这里不能谈它,因为我们必须谈微分方程的重大论题.

我们甚至不可能谈庞加莱根据前人的观点在微分方程的研究上所取得的进步.但是,在沿着这些道路继续前进的同时,他又开辟了一条崭新的富有成果的道路.这条新道路放弃了复数域,而探索满足方程的实曲线的一切可能的形式.

在这里,在数学史的其他问题上曾经发生过的一种情况又出现了.当问题涉及一个代数方程的解时,第一流的代数学家,直到 18 世纪最后四分之一世

纪,都在孤立地讨论方程的一个根,但他们最终在这条道路上停下来了.当对这些根的集合以及根之间可能存在的关系进行探讨时,关于这个问题所进行的真正富有成果的研究才全面展开,庞加莱教导我们,我们在研究微分方程的时候,也要遵循这同一个原则.

人们对满足假定从一阶开始的同一个微分方程的各种曲线能说些什么呢?简单地讲就是,它们中间任两条除了在奇点外永远不可能相切.这个乍看起来脆弱的基础作为庞加莱的出发点是足够了.

然而,他在这里又加进了一个被他的所有前辈所忽视了的基本意见.一个乍一看来十分幼稚的意见被黎曼遗漏了,而这个表面上看来无知的意见却在科学中非常卓越地发挥了根本性作用.如果说柯西,在其发表的第一篇论文(1813年)中,不曾有这方面的论述的话,那么,黎曼在40年之后也没有得到完成代数函数论的荣誉,当轮到黎曼认识并运用这个原则去完成这一理论的时候,他并不明白此原则不仅仅只是适用于他刚刚研究的问题.庞加莱则能够指出此原则是微分方程研究中必不可少的,并且多亏了他,才探讨了这些方程所定义的积分曲线的一切可能的形式.

黎曼(Riemann, Georg Friedrich Bernhard, 1826—1866),德国数学家.

人们可以根据初等可积方程的简单情况所提供的例子来推想这些曲线可能存在的形式.这样作,会出现差错,凡是关心科学方法的人,大概都会考虑这一点.这是我们在某些特殊情况下知道怎么处理的一个问题.我们根据这些特殊情况下已知解所显示的特点,去推论我们感到类似的其他问题.我们不大可能不这样做.但是,不要忘记,它们彼此之间必然会存在着深刻差别,虽然这些差别我们看不出来,但我们有能力解决一些问题而无能力解决另一些问题这个事实本身大概会提醒我们.

一阶微分方程的实际情况就是这样.如果一阶微分方程像初等情形中那样,可以借助于积分来处理,亦即使坐标的某一个函数与一个常量相等,那么,它就不可能得出庞加莱所明确指出的那样的图形.

但是,一阶微分方程的研究仅仅是第一个阶段.二阶方程的研究使庞加莱接近于他的主要课题.他用全部的现代分析,不断地注视着这个课题:天体力学的微分方程,这些方程决定着各行星的运动,这种运动不仅受到太阳引力的影响,而且也受到它们相互引力的影响.

谈不上在这样一个初等意义的方程组上求积的问题.为此,必须掌握足够数目的积分.而实际人们所掌握的还差得很远,因此庞加莱指出,必须放弃给十个经典积分再增加任何一个积分的想法.

相反地,另一方面,一个头等重要的事实支配了他的研究:周期解的存在.如果一颗行星(例如地球)是面对着太阳的唯一行星并且只受到太阳的引力,那





么它就会如托勒密甚至开普勒所想象的那样运动,就是说,在一年的岁终,它正好又回到它原来出发的位置.并且,在同一个运动的推动下,开始了与第一年完全一样的第二年的运行,而且这种运行是周期性的,实际情况并不是这样,因为地球还受到其他行星的引力.此外,这些行星之间也互相吸引,干扰彼此的运动.甚至在这些新条件下,太阳系——简化为由最简单情形开始的三体问题——是否有可能存在周期运动呢?这就是庞加莱不断向自己提出的问题,并且,从我们谈到过的他的特别富有成果的早年研究开始,直到死神窥伺着他的时刻,直到其悲剧性的序言提及他对自己的健康非常担心这最后一篇论文为止都是这样.

托勒密(Ptolemy,约100—170),希腊天文学家、数学家.

开普勒(Kepler, Johannes, 1571—1630),德国数学家.

他这样发现的这些周期解向人们指出了一条可以深入到一直被人认为是高不可攀之处的唯一途径.

由这种观点来看,二阶微分方程较简单的情形,即在通常的空间中确定一条曲线的情形,已经是典型化了的.如果一条这样的曲线是封闭的,那么,相邻的那些曲线解大概也将是封闭的,而且其中起点已定的任何一条曲线,都周期性地重新经过同一个点的邻域.但是,这些连续弧所呈现的排列——在某些情况下庞加莱所创立的方法允许十分严谨地进行讨论的排列——一方面表明,当人们在将来和过去同时观察研究这些运动的时候,是何种奇异得几乎不可理解的排列能够呈现这些运动;另一方面表明,人们在天体力学的经典方法中所遇到的困难同事物的本质何等相关.

当他在相反的意义上接触到这些文字性结论,而这些结论说明人们指望的这些进展相对说来是靠不住的时候,庞加莱就另外开辟了一条容许在坚实的基础上继续前进的道路.他指出了其作用的这个概念与这些同样的积分相类似,我们知道,进一步地研究这些积分是徒劳无益的.不变积分的概念再一次表明,不是对一个孤立的解,而是对全体解或至少一部分解进行考察时,是多么有用.这是一个积分,也就是一个根据微分方程保持不变的量,而且,这是一个建立在解集合上的共同积分.就是这个概念,在庞加莱的手中特别得出了我们关于对太阳系稳定性这个如此根本性的问题的最基本的结论.

名为 *Les Méthodes nouvelles de la Mécanique céleste*(《天体力学的新方法》)的著名著作就是专门论述这些新的有效的方法的.不仅在天文学上,而且在其他许多力学问题上,这些方法都发挥着几益明显的作用.

我们仅仅很不完全地谈到在微分方程的广阔领域中应归功于庞加莱的丰功伟绩.但是,从科学研究偏微分方程的时刻起,由物理应用提出的另一些问题就吸引了科学.这些问题,由于新出现的某些复杂情况,人们一度可能认为是和上面的问题相似,但实际上则完全不同,其典型已由著名的狄利克雷问题提供了.

狄利克雷(Dirichlet, Pet Gustav Lejeune 1805—1859),德国数学家.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

在这个引人入胜的问题上，花费了多少心血啊！在为研究这个问题而提出的许多方法之中，庞加莱又增添了一种新的方法，扫描的方法．这个增添是否具有重大意义呢？这项理论后来的进程做出了响亮的回答．从扫描方法出现起，关于这个问题的所有研究工作都靠这个方法．从那时起，这个方法就一直支配着我们关于这个问题的所有见解．

许多重大的发现，我是应该提一提的，可我在这里不可能谈了，虽然它们也构成此后进步的必不可少的基础．当死神把庞加莱从我们中间夺走时，科学还没有权利把那么多另外的成果，那么多基本观点期待于他．

他给我们留下的财富，我只能简短地不充分地给一个大概的回顾，半个多世纪以来，这些财富是我们科学进步的基础，并且在许多点上也是人类精神主要进步的基础．

2 数学的未来①

为了预见数学的未来，正确的方法是研究它的历史和现状．

在某种意义上，这不正是我们数学家的专业程序吗？我们习惯于外插法，这是一种从过去和现在推测未来的方法，因为我们充分地了解这种方法的含义，所以关于它给予我们的结果的有效范围，我们不会冒受骗的危险．

迄今我们已经有了不幸的预言家．他们轻率地重申，所有能够解答的问题都被解决了，除了拾遗之外，没有什么可作的了．幸好，过去的情况使我们消除了疑虑．有人往往以为，所有问题都被解决了，或者至少已列出了一切容许的解的清单．可是，解这个词的意义扩大了，不可解的问题变得最使大家感兴趣，未曾料到的其他问题呈现在眼前．对于希腊人来说，好解就是只使用直尺和圆规的解；后来，它变成用求根法得到的解，接着它又变成只利用代数函数和对数函数的解．这样一来，悲观主义者发现他们自己总是被挫败、总是被迫退却，因此我现在认为不再有悲观主义者了．

所以，我的注意力并不是反对他们，因为他们已经消亡了，我完全知道，数学将继续发展，但是问题在于它怎么发展、向什么方向发展？你将回答："在各个方向发展．"这部分地正确；如果它完全正确，那可有点骇人听闻了．我们的丰富资料不久便会成为累赘，资料的积累也会造成一堆费解的大杂烩，犹如无知的人面对未知的真理那样莫名其妙．

① 原载自：庞加莱．科学的价值[M]．李醒民，译．北京：光明日报出版社，1988：359-373．





历史学家、物理学家甚至都必须在事实中做出选择,科学家的头脑只能顾及宇宙之一隅,永远也不能囊括整个宇宙,以至于在自然界提供的不可胜数的事实中,一些将被忽略,另一些则被保留下来.

不用说,在数学中正是这样,几何学家已不能迅速地把握所有呈现在他面前的杂乱的事实,更有甚者,正是他——我几乎要说他的任性——创造这些事实.他把各个元素收集在一起构造全新的组合;一般说来,自然并没有把预先准备好的组合给予他.

毫无疑义,有时会发生这种情况:数学家着手解决问题是为了满足物理学的需要;物理学家或工程师请求他计算某些应用方面的数值.难道能够说,我们几何学家应当仅限于听候命令,而不因我们自己的欢娱来经营科学,只是力图使我们自己迁就我们庇护人的需求吗?假如数学除了帮助那些研究自然的人而外没有其他目的,那我们就只好听候命令了.这种看问题的方式合理吗?这绝不合理;如果我们不为科学而经营精密科学,那我们就不可能创造出数学工具,待到物理学家提出请求的那一天,我们就会无能为力.

物理学家研究一种现象,也不是要等到物质生活的某种急迫需要使它成为他们必不可少的东西;他们是对的.假使 18 世纪的科学家因为电在他们的眼中只是好奇的玩意儿而没有实际利益,因此忽略电的研究,那么在 20 世纪,我们就不可能有电报,也不可能有电化学、电技术.所以,不得不进行选择的物理学家并没有仅仅以实用性来指导他们的选择.可是,他们怎样在自然事实之间选择呢?使他们感兴趣的事实是能够导致发现规律的事实,这些事实因而类似于许多其他在我们看来似乎不是孤立的,而是与另外的东西聚集在一起的事实.

孤立的事实吸引着大家的眼睛,吸引着外行人的眼睛和科学家的眼睛.但是,唯有名副其实的物理学家才知道,如何观察把许多事实联系在一起的结合物是哪一个,这些事实的类似是深刻的,却是隐蔽的.牛顿的苹果故事恐怕不是真实的,而是象征性的,不过,让我们把它当做真实的谈一谈吧.我们必须认为,在牛顿之前,好多人都看见过苹果落地;没有一个人知道从中得出什么结论.假如没有能够在其背后隐藏着某种东西的事实中做出分辨和选择,以及识别什么正在隐藏着的精神,假如没有在未加工的事实中察觉事实精髓的精神,事实也是毫无成果的.

牛顿(Newton Isaae, 1643—1727),英国数学家、物理学家、天文学家.

我们在数学中正好发现同样的东西.从我们正在处理的各种各样的元素中,我们能够得到无数个不减的组合,但是,这些组合中的一个倘若是孤立的,则其毫无价值可言.我们常常含辛茹苦地构造它,但是它却没有效用,也许至多不过是为初等教育提供练习而已.当这个组合在一类相似的组合中找到了位置时,当我们注意到这种类似时,它就完全是另外一个样子了.我们就不再是面对一个事实,而是面对一个定律.在那一天,真正的发现者将不是耐心地建造某些

组合的工匠;真正的发现者将是揭示它们的亲缘关系的人.前者看到的只是未加工的事实,只有后者才能察觉到事实的精髓.往往为了确定这种亲缘关系,足以使他创造出新名词,这个名词是有创造力的.科学史向我们提供了大量大家都熟知的例子.

著名的维也纳哲学家马赫(Mach)曾经说过,科学的作用在于产生思维经济,正像机器产生劳力经济一样.这是十分正确的.原始人用他的手指或借助于石子来计算.在给儿童教乘法表时,我们使他们以后节省了用一堆石子进行计算的辛劳.某人已经发现,用石子或其他东西计算,6 乘 7 等于 42,他特意把这个结果记录下来,因此我们不需要重复它了.他没有白费他的时间,即使他是为消遣而计算的:他作运算只花了两分钟,如果十亿个人在他之后重复作这个运算,那总共就要花费二十亿分钟时间.

于是,事实的重要性用它产生的效益来衡量,也就是说,用它容许我们节省的思维数量来衡量.

在物理学中,具有最大效益的事实是进入十分普遍的定律中的事实,这些事实能够使我们根据定律预见大量的其他事实,在数学中情况正是如此.设想我从事一项复杂的运算,费力地达到了一个结果:如果我由此还不能预见其他类似运算的结果,还不能可靠地指导运算以避免人们在首次尝试中不得不进行的摸索,那么我就没有补偿我的辛劳.另一方面,如果这些摸索本身最终向我揭示出刚刚处理的问题类似于更为广泛一类的其他问题,如果它们一举向我表明这些问题的相似和差异,一句话,如果它们使我察觉到推广的可能性,那么我就没有白费我的时间.因此,这不是我已经赢得的一个新结果,而是一种新的能力.

首先想到的简单例子是代数公式的例子,当我们最后用数字代替字母时,这些公式便把一种类型的数值问题的答案给予我们.多亏它,一次代数运算就使我们省去不断重新开始新的数值计算的辛苦.但是,这只是一个粗糙的例子,我们大家都知道,还有不能用公式表示的、更为珍贵的类似性.

既然在把早就已知的而迄今依然分离的、似乎相互陌生的元素统一起来时,新结果突然在表面上由不一致统治的地方引入秩序,那么它就是有价值的.于是,它容许我们一眼看到这些元素中的每一个以及它在集合中的位置.这个新事实不仅仅因其自身而珍贵,而且唯有它才能使它所结合的一切旧事实具有价值.我们的精神像我们的感官一样,也是软弱的,如果世界的复杂性不是和谐的,它就会在这种复杂性中茫然无措,它像一个眼睛近视的人一样,只能看到细枝末节,在审查下一个枝节前,它又会被迫忘掉先前的每一个枝节,因为它不能囊括整体.唯一值得我们注意的事实是那些把秩序引入到这种复杂性中去的事实,从而是使我们的精神可以理解的事实.





数学家把重大的意义与他们的方法和他们的结果的雅致联系起来.这不是纯粹的浅薄涉猎.在解中、证明中给我们以雅致感的实际上是什么呢？是各部分的和谐,是它们的对称、它们的巧妙平衡；一句话,雅致感是所有引入秩序的东西,是所有给出统一、容许我们清楚地观察和一举理解整体和细节的东西.可是,这正好就是产生重大结果的东西,事实上,我们越是清楚地、越是一目了然地观察这个集合,我们就越是彻底地察觉到它与其他邻近对象的类似性,从而我们就有更多的机会推测可能的推广.在意外地聚合了我们通常没有汇集到一起的对象时,雅致可以产生意想不到的感觉,在这里,它再次是富有成果的,因为它这样便向我们揭示出以前没有辨认出的亲缘关系.甚至当它仅仅起因于方法的简单性和提出的问题的复杂性之间的悬殊差别时,它也是富有成效的；于是,它促使我们想起这种悬殊差别的原因,而且每每促使我们看到,偶然性并不是原因,它必定能在某个未曾料到的定律中找到.简言之,数学雅致感仅仅是由于解适应于我们精神的需要而引起的满足,这个解之所以能够成为我们的工具,正是因为这种适应.因此,这种审美的满足与思维经济密切相关.我又一次想到埃雷克泰米的优美的女像柱的比喻,但是我没有必要过于经常地利用它.

正是由于同样的理由,当相当冗长的运算导致出某一简单的引人注目的结果时,只要我们还未证明,我们即使不能预见完整的结果,至少应该能够预见它的大多数特性,那我们就不会心满意足.为什么？这个运算似乎把我们想要知道的一切都告诉给我们,究竟是什么东西妨碍我们以此为满足呢？这是因为,在类似的情况中,冗长的运算不再起作用了,与此相反,关于能使我们预见的推理——它常常有一半是直觉的——却不是这样.由于这种推理简短,我们一瞥就可见到它的所有部分,以致我们直接觉察到,为了使它适应于能够出现的同一本性的问题,我们必须改变什么.因此,它能够使我们预见这些问题的解是否将是简单的,它至少能够向我们表明是否值得从事这一运算.

我们刚才所说的一切足以表明,企图用任何机械程序来取代数学家的自由的首创精神,将是多么愚蠢啊.为了得到具有真正价值的结果,刻苦地进行运算,或者拥有整理事物的机械,都是不够的；值得花时间追求的不只是秩序,而是未曾料到的秩序.机械可以抓住未加工的事实,却总是抓不住事实的精髓.

自 19 世纪中期以来,数学家越来越渴望得到绝对的严格性；他们是对的,这种倾向将越来越受到重视.在数学中,严格性不是一切,但是没有严格性便没有一切.不是严格的证明毫无价值可言.我想没有人对这个真理持有异议.但是,如果过分照字面来理解严格性,我们可能会得出结论,例如在 1820 年之前,还不存在数学；这显然是过分了；当时的几何学自发地理解了我们现在用冗长的论述解释的东西.这并不意味着他们根本没有看到严格性；而是他们太迅速地越过了它,要明确地领会它,就必须费心把它讲一讲.

但是，总是需要这么多的次数来讲它；在所有其他人之前第一个强调精密性的人，把我们可以力图仿效的论据给予我们；可是，如果将来的证明都建立在这个模型的基础上，那么数学论文就会十分冗长；我担心冗长起来，不仅因为我不赞成书刊塞满图书馆，而且因为我担心这样冗长下去，我们的证明就可能失去和谐的外观，我刚才已说明了和谐的用处.

思维经济是我们应该追求的目标，因此只提供仿效的模型还是不够的.需要使我们之后的人能够省却这些模型，不去重复已经做出的论据，而用几句话概括它.而且，这一点时常被达到.例如，有一种到处都可找到的推理形式，而且处处相似.它们是完全精密的，但却颇为冗长.于是有一天，有人想到“收敛的一致性”这个用语，这个用语使这些论据变得不需要了；我们不再必须重复它们了，因为它们可以被理解.这样一来，那些克服过于繁琐的困难的人就给我们双重的帮助：他们首先告诉我们在紧急时像他们那样去作，但是尤其是，他们能够使我们在不牺牲精密性的条件下，尽可能经常地避免像他们那样去作.

我们刚才通过一个例子已经看到名词在数学中的重要性，不过还可以引用许多其他例子.正如马赫所说，人们很难相信，一个精选的名词就能使思维有多么经济.也许我在某处已经说过，数学是把同一名称给予不同事物的艺术.可以说，把这些在内容上不同而在形式上可以相似的事物纳入同一模式中是恰当的.当选好语言时，我们不胜惊讶地发现，对某一对象所作的论证可直接用于许多新对象；这里什么也没有改变，甚至连名词也没有改，因为名称已变成相同的了.

一个精选的名词通常足以消除用旧方法陈述的法则所遇到的例外，这就是为什么我们创造了负数、虚数、无穷远点等.我们一定不要忘记，例外是有害的，因为它掩盖着定律.

好了，这是我们用以辨认产生巨大结果的事实的特征之一.多产的事实是容许这些巧妙的语言革新的事实.因此，未加工的事实往往不会使人有多大兴趣；我们可以多次指出它，但对科学并没有提供多大帮助.只有当比较有见地的思想家觉察到它所代表的关系，并用名词表征它时，它才会获得价值.

此外，物理学家的做法正好相同.他们发明了“能”这个词，这个词格外富有成效，因为它通过消除例外而创造了定律，由于它把同一名称给予内容不同而形式相似的事物.

在具有最幸运的影响的名词中，我可以挑选出“群”和“不变量”.它们使我们看到许多数学推理的实质；它们向我们表明，在多少情况下，以往的数学家是在不明其义的情况下考虑群的，他们是怎样突然发现它们不知何故这么接近，他们原以为它们相距甚远呢！





今天,我们可以说,他们处理的是同构群.我们现在知道,在一个群中,内容几乎没有什么令人感兴趣之处,唯有形式才有考虑的价值,当我们了解一个群时,我们从而也就了解所有的同构群;由于“群”和“同构”这些名词把微妙的法则浓缩在几个符号内,并使所有的精神迅速地通晓它,因此转变是直接的,能够以充分的思维经济去完成.此外,群的观念归属于变换的观念.我们为什么要把这样的价值放在新变换的发明上呢? 因为从一个定理,能使我们得到十个或二十个定理;这与在紧靠整数的右边加一个零具有同样的价值.

这一切就是迄今决定数学进展方向的东西,将来肯定还会完全一样地起决定作用.为此,所提出的问题的性质同样起一份作用.我们不能忘记,我们的目的应当是什么.按照我的观点,这个目的是双重的.我们的科学与哲学和物理学二者相毗邻,我们为我们的两个邻居而工作,因此,我们总是看到,而且还将继续看到数学家在两个相反的方向上前进.

另一方面,数学科学必须反省自身,由于反省自身就是反省创造它的人类精神,因而是有用的,因为数学科学是人类精神从外界借取的东西最少的创造物之一,所以它就更加有用了.这就是为什么某些数学推测是有用的,例如专门研究公设、非寻常几何、特殊函数的推测.这些推测距通常的概念以及距自然界和应用越远,它们就越加充分地向我们表明,当人类精神越来越多地摆脱外部世界的羁绊时,它能够创造出什么东西,因此它们就越加充分地让我们在本质上了解人类精神.

但是,我们必须指挥我们的主力部队向其他方面前进,即向自然界方面前进.

在这里,我们遇到了物理学家和工程师,他们对我们说:“请给我积分这个微分方程吧;我可能在一周后需要它,由于到那时我要完成一项工程.”我们回答说:“这个方程不能归入一种可积类型,你也知道可以积分的方程并不多.”“是的,我知道;但是到那时你有什么作用呢?”通常相互谅解也就够了;工程师实际上不需要无限项的积分;他需要知道积分函数总的概貌,或者说,他只要求实际上能够从这个积分推演出来的某一个数,倘若这个积分已知的话.通常它不是已知的,但是没有它我们也能计算出这个数,只要我们确切地知道工程师需要什么数以及要达到什么近似程度就可以了.

从前,只有当一个方程的解能够借助于有限数目的已知函数表示时,人们才认为解了方程,但是,这种可能性百中难得其一.我们经常能够做的,或者恰当地讲,我们应该经常力图去做的,可以说是定性地解决问题;也就是说,力图去了解表示未知函数的曲线的一般形状.

依然要寻找问题的定量的解;可是,如果未知数不能用有限的运算确定,那

莱布尼兹(Leibniz, Gottfried Wilhelm, 1646—1716),德国自然科学家、数学家、哲学家.

么它总是可以用能使我们计算它的有限收敛级数来表示.能够认为这是它的真实解吗?我们听说,牛顿曾给莱布尼兹寄了一个字谜,其内容大致是 aaaaabbbeeeeii 等.莱布尼兹当然一点也不理解它;但是,我们掌握这个字谜的秘诀,了解它的意思,把它译成现代词语就是"我能够积分一切微分方程",而我们却被诱使说,牛顿要不就是十分幸运,要不就是有奇怪的错觉.牛顿只是想说,他能够形成(用待定系数法)一个形式上满足所提出的方程的幂级数.

这样的解今天不会使我们满意,其理由有二:因为收敛太慢,因为相互紧随的项不服从任何规律.相反地,Θ 级数在我们看来好像是完美无缺的,首先因为它收敛很快(这有利于希望尽可能快地得到一个数的实用者),其次因为我们一瞥即见项的规律(这可以满足理论家的审美要求).

但是,这样一来,就不再有可解的问题和不可解的问题;有的只是或多或少可以解决的问题,或者它们可以通过大体上迅速收敛的级数来解,或者它们由大体上和谐的定律来支配.不过,往往发生这种情况:不完美的解把我们引向比较完美的解.有时,级数收敛过慢,以至于计算无法实际进行,我们仅仅得以证明问题的可能性.

因此,这位工程师觉得这是一种嘲弄,这恰恰由于它没有帮助他在规定的日期完成他的工作.他不想了解它是否将会有益于 22 世纪的工程师们.但是,至于我们,我们却不这么认为,能为我们后代省却一天工作,有时也比为我们同代人节约一个小时更为使我们满意.

可以说,有时通过从经验上摸索,我们达到一个充分收敛的公式.工程师说:"你们还需要什么?"可是,不管怎样,我们仍不满足;我们希望预见那个收敛,为什么?因为只要我们一次知道怎样预见它,我们就会知道下一次怎样预见它.我们成功了;如果我们不能再次有效地预期这样做,那么成功也不过是小事一桩而已.

随着科学的发展,对它作整体的理解也变得更加困难;于是,我们企图把它分成小块,而满足于这些小块之一:换句话说,企图使它专门化.如果我们在这条道路上继续走下去,那就会为科学的进步设置严重的障碍.正如我们所说,科学进步正是由于它的各部分之间未曾料到的结合引起的.过分专门化便会妨碍这些结合.希望像海德堡会议和罗马会议这样的会议,通过使我们彼此之间接触,向我们打开邻近领域的视野,促使我们把邻近领域与我们自己的领域加以比较,以便探寻我们自己的小群落之外的东西;因此,这样的会议将是对刚才所提到的危险的最好补救方法.

然而,我在一般原则上拖延的时间太长了;现在是逐一详述的时候了.

让我们分别审查一下各门特殊学科,它们联合起来构成了数学:让我们看看,每门学科完成了什么,它向哪里发展,我们对它可以有什么希望.如果原先





的观点是正确的,那么我们就会看到,过去的最大进展发生在这些学科中的两个结合之时,发生在我们开始意识到它们形式的类似性而不管它们内容的差别之时,发生在它们相互之间如此模仿以至一个获胜而另一个也受益之时.与此同时,我们可以在同类的结合中预见未来的进步.

2.1 算术

与代数和分析相比,算术的进步很慢,容易看到其中的原因.连续性的感觉是一种宝贵的指导,但是算术家却缺少它,每一个整数都与其他整数相分离——也就是说它具有自己的独立存在性.它们中的每一个都是一种例外,这就是在数论中普通定理比较稀少的原因,这也是存在的定理隐藏得比较多,而且比较长期地使研究人员为难的原因.

如果说算术落后于代数和分析,为此最有效的做法是,力图使算术仿效这两门学科,以便从它们的进展中获得好处.因此,算术家应当把与代数的类似作为向导.这些类似是大量的,在许多情况中,即使还没有充分仔细地研究它们是否可以利用,但至少长期以来已预见到它们,甚至两门学科的语言表明,人们已清楚地认识到它们.我们这样谈论超越数,我们这样说明超越数的未来分类是超越函数分类的模型,我们迄今还没有十分明确地看到如何从一种分类过渡到另一种分类;但是,假如人们已认识到它,那么它已经被完成了,它已不再是将来的工作了.

我想起的第一个例子是同余理论,在其中可以找到与代数方程理论完全的平行性.的确,我们将会成功地完成这种平行性,例如,这种平行性必须在代数曲线理论和两个变量的同余理论之间成立.而且,当要解决与几个变量的同余有关的问题时,这将是通向解决许多不定分析问题的第一步.

2.2 代数

代数方程理论还将长期地引起几何学家的注意,人们可以着手研究的方面是很多的、很不同的.

我们不必认为代数走到了尽头,因为它给我们以形成所有可能组合的规则;依然要寻找满足某种条件的有趣的组合.这样一来将形成一种不定分析,其中的未知数将不再是整数,而是多项式.这时,代数将仿效算术,以至于整数与具有任意系数的整多项式或具有整系数的整多项式可以类比.

2.3 几何学

看起来,好像几何学包含的无非是在代数或分析中已经包括的东西,几何学事实只不过是用另一种语言表达的代数事实或分析事实.因此,可以认为,在

我们考察之后,没有多少特别与几何学有关的东西供我们谈论了.这也许是没有明确认识到精心构造的语言的重要性,不理解借助于表示这些事物的方法以及分类它们的方法把什么添加到事物本身之中.

首先,几何学的考虑导致我们提出新问题;如果你乐意的话,这些问题可以是解析问题,但是在解析方面,我们永远也提不出这样的问题.不过,解析因这些问题受益,正如它因满足物理学的需要而必须解决的问题受益一样.

几何学的巨大优点在于,感觉能够帮助思维,有助于发现前进的道路,许多人都偏爱把解析问题化为几何学形式.不幸的是,我们的感觉不能把我们带得很远,当我们想超越经典的三维时,感觉就无法帮助我们.这难道意味着超过感觉希望把我们束缚于其内的有限领域,我们只能依靠纯粹解析,所有大于三维的几何学都是徒劳的和无目的的吗? 前一辈的大师们可能回答"是";今天,我们如此熟悉这个概念,以至于我们甚至能在大学课程中谈及它了,而不会引起过多的惊讶.

但是,它有什么用处呢? 这很容易看到:首先,它给我们以十分方便的术语,这种术语简洁地表达了通常解析语言用冗长的词语才能讲清的东西.而且,这种语言使我们用同一名称称谓相似的事物,使我们突出类似性,让我们永远不再忘记它.因此,这种语言还能使我们在多维空间中找到我们的道路,并使我们联想起视觉空间,对我们来说,多维空间太大了,我们不能看到它,视觉空间无疑只是它的不完善的影像,但不管怎样总是一种影像.与前面所有的例子一样,这里又一次出现了与简单事物的类似性,这使我们能够理解复杂事物.

这种大于三维的几何学不是简单的解析几何学,它不是纯粹定量的,而是定性的,正是在这方面,它变得尤其有趣.有一种学科叫拓扑学,它把研究图形的不同元素的位置关系作为它的对象,而不管各元素的大小.这种几何学是纯粹定性的;即使图形不是严格的,是一个孩子粗略地描画的,其定理却依然为真.我们也可以创造出大于三维的拓扑学.拓扑学的重要性是巨大的,但也不能强调得太过分了;拓扑学的主要创始人之一黎曼从中得到的好处足以证明这一点.我们必须得到它在多维空间中的完整结构;我们因而将有一种工具,使我们实际上能在多维空间内观察,以弥补我们的感觉.

假如只讲解析语言,那么拓扑学问题也许还不会产生;或者确切地讲,我弄错了,这些问题确实出现了,由于它们的解答对于一大堆解析问题是必不可少的,但是它们是一个接一个地单独来到的,我们未能察觉到它们共同的结合物.

康托(Conner, Pierre Eudide Jr., 1932—),美国数学家,弗吉尼亚大学教授.

2.4 康托主义

上面我已说过,我们需要继续回到数学的第一原理,还说这对研究人类精神有什么好处.这种需要唤起了两种努力,它们在最近的数学编年史上占据了





十分重要的位置.其一是康托主义,它对数学的贡献是尽人皆知的.康托把一种新的考虑数学无穷的方法引入科学.康托主义的特征之一在于,它不是通过建立越来越复杂的构造上升到一般,不是通过构造来定义的,正如学究们所说的,它从最高类出发,只通过最近类和种差来定义.它在一些人的思想上引起的极端厌恶便由此而来,例如埃尔米特即是其中之一,他特别偏爱的观念就是把数学科学和自然科学加以对照.就我们大多数人而言,这些偏见已经烟消云散了,但是却出现了这样的情况:我们意外地碰到了某些悖论,即某些表面上的矛盾,它们会使爱利亚人芝诺和迈加拉学派(the School of Megara)感到高兴.因此,每一个人都必须寻求补救办法.就我个人——不仅是我一个人——而言,我认为,重要的事情永远不是引入不能用有限的词完善定义的实体.不管采纳什么治疗方案,我们都可以指望像请来的医生那样高兴,因为他是仿效绝妙的病理学方面的案例来诊治的.

芝诺(Zeno of Elea,约前490—前430),希腊哲学家.

2.5 公设的探讨

另一方面,人们曾努力列举或多或少隐藏的公理和公设,把它们作为各种数学理论的基础.希尔伯特教授获得了最辉煌的成果.乍看起来,这个领域似乎是很有限的,当花不了多长时间把目录编好后,就不会有什么事可做了.但是,当我们清点了所有的东西后,就将有许多种分类这一切的方法;一个健全的图书馆总可以找到某些要做的事情,每一种新分类法都会对哲学家有所启发.

希尔伯特(Hilbert, David,1862—1943),德国数学家.

在这里,我要结束这一考察了,我不会梦想使它完善.我认为,这些例子足以表明,数学科学通过什么机制在过去取得了进步,它将来必须在什么方向上进展.

3 数学的创造①

数学创造的发生是一个使心理学家强烈感兴趣的问题.它是一种活动,在这种活动中,人类精神似乎从外部世界所取走的东西最少,在这种活动中,人类精神起着作用,或者似乎只是自行起作用和按照自己的意志起作用,以至于在研究几何学思维的步骤时,我们可以期望达到人类精神的最本质的东西.

长期以来,人们已意识到这一点,数月之前,由莱桑(Laisant)和费尔(Fehr)编辑的《数学教学》杂志开始调查各个数学家的思想习惯和工作方法.当那项调查的结果发表时,我已写完了我的这篇文章的主要轮廓,因此我不可能利用他

① 原载自:庞加莱.科学的价值[M].李醒民,译.北京:光明日报出版社,1988:374-386.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

们的调查结果，我只想说明，大多数证据证实了我的结论；我没有说全部证据，因为当诉诸全体投票时，不可能希望会取得完全一致.

最初的事实会使我们感到惊奇，或者更确切地讲，如果我们还不这样习惯它，它就会使我们感到惊奇.有人不理解数学，这是怎么发生的呢？既然数学有求助于所有正常思想都能接受的逻辑规则，既然数学的证据建立在对一切人都是共同的原理的基础上，既然没有一个不发疯的人会否认这一点，那么在这里为何出现如此之多的不开化的人呢？

并非每一个人都能够发明，这决不是难以理解的.并非每一个人都能够记住一次学到的证明，这也可以略而不提.但是，当把数学推理加以解释之后，并非每一个人都能够理解它，我们想想这件事，似乎是十分奇怪的.可是，大多数人只能够相当吃力地仿效这一推理：这是不可否认的，中学教员的经验确实不能否认这一点.

进而要问：在数学中为何也可能出错？健全的头脑不应当有逻辑谬误，可是有一些十分敏锐的人，他们对于发生在日常生活各种活动中的简短推理未犯错误，但是却不能毫无错误地仿效或重复数学证明，这些证明虽则较长，但毕竟只是完全类似于他们容易做出的简短推理的累积.我们还需要补充说数学家本人也不是一贯正确的吗？

答案对我来说是显而易见的.设想一个长系列的三段论，第一个的结论是下一个结论的前提：我们将能够理解这些三段论的每一个，在从前提向结论的过渡中，我们没有处于受骗的危险之中.但是，我们在某一时刻遇到一个命题作为一个三段论的结论，在过去一些时间后，我们又偶然地重新遇到它却作为另一个三段论的前提，在这两个时刻之间，这个链条的几个环节将显示出来；这样一来，可能会发生下述情况：我们忘掉了这个命题，或者还要糟糕，我们忘掉了它的意义.因此，我们可以用一个稍微不同的命题代替它，或者在保持同一阐述时，我们却赋予它以稍微不同的意义，我们面临的错误就是由此而来.

数学家常常使用法则.他当然以证明这个法则开始；当这个证明在他的记忆中还很清晰之时，他完全理解它的意义和它的关系，他没有陷入改变它的危险中.可是后来，他相信他的记忆，并进而仅仅以机械的方式应用它；不过，如果他的记忆模糊不清了，他就会统统把它用错.举一个简单的例子，这就像我们之所以有时在运算中出差错，是因为我们忘记了我们的乘法表.

照此看来，数学的特殊能力仅仅是由于十分可靠的记忆或惊人的注意力.它是一种类似于玩惠斯特牌的人的能力，玩牌人记着所玩的牌；或者再进一步，它类似于棋手的能力，棋手能够想象为数众多的组合，并把它们记在心里.每一个高明的数学家都应该是高明的棋手，反过来也是这样；同样地，他也应该是一位高明的计算家.不用说，这有时会出现；例如，高斯同时是天才的几何学家和





十分早慧的、精确的计算家.

但是也有例外;或者确切地讲,我弄错了;在例外不多于常规的情况下,我不能把它们叫做例外.相反地,正是高斯,才是一个例外.至于我自己,我必须承认,即使让我作加法,我也绝对不可能不出错.按同样的方式,我无非是一位蹩脚的棋手;我察觉到,采用某种走法,我会陷于某种危险;我审查了几种其他走法,因另外的理由放弃了它们,而最后我走了第一次考察的步子,其间我忘记了我已经预见到的危险.

换句话说,我的记性不是不好,但它不足以使我成为一位高明的棋手.在一段困难的数学推理中,最高明的棋手也会晕头转向,可是我为什么不会失败呢?显然,这是因为一般的推理步骤引导着我.数学证明不是演绎推理的简单并列,它是按某种次序安置演绎推理,这些元素安置的顺序比元素本身更为重要.如果我具有这种次序的感觉,也可以说这种次序的直觉,以便一眼就觉察到作为一个整体的推理,那么我已无须害怕我忘记这些元素之一,因为它们之中每一个都在排列中得到它的指定位置,而且不要我本人费心思记忆.

于是,在我看来,在重复已学到手的推理时,好像我能够发明它.这往往只是一种错觉;但是,即便如此,即便我本人没有创造它的天赋,就我重复它而言,我本人也是重新发明它.

我们知道,数学次序的这种感觉、这种直觉,使我们推测隐蔽的和谐与关系,但它并不是每一个人都具有的.有些人或者没有这种如此难以定义的微妙的感觉,或者没有超常的记忆力和注意力,因此他们绝对不可能理解较高级的数学.这种人是多数,另一些人仅略有这种感觉,但他们具有非同寻常的记忆力和高度的注意力这样的天赋.他们将一个接一个地记住各种细节;他们能够理解数学,有时也能应用,但他们不能创造.最后,还有一些人或多或少地具有所提到的特殊的直觉,因此,即使他们的记忆力毫无非同寻常之处,他们却不仅能够理解数学,而且可以成为创造者,并试图做出发明,其成功之大小取决于这种直觉在他们身上发展的程度之大小.

数学创造实际上是什么呢?它并不在于用已知的数学实体做出新的组合.任何一个人都会作这种组合,但这样做出的组合在数目上是无限的,它们中的大多数完全没有用处.创造恰恰在于不做无用的组合,而做有用的、为数极少的组合.发明就是识别、选择.

我以前已解释过如何进行这种选择;值得加以研究的数学事实是这样一些事实,通过它们与其他事实的类比,能够导致我们了解数学定律,正像实验事实能够导致我们了解物理学定律一样.它们是向我们揭示其他事实之间意外关系的事实,我们虽然早就知道其他事实,但却错误地认为它们彼此之间是毫无联系的.

在所选择出来的组合中,最富有成果的组合常常是从相距很远的领域取来的元素所形成的组合.我的意思并不是说,把尽可能相异的对象收集到一起就足以做出发明;这样形成的大多数组合都毫无成果.但是,它们之中的某些极稀有的组合却是最富有成果的.

我说过,发明就是选择,但是,这个词也许不完全严格.它使人们想起采购员,在他面前陈放着大量的货样,他逐个审视它们,以便做出选择.这里的货样多得不可胜数,他的整个一生也难以把它们审查完.事情的实际状况并不是这样.无结果的组合甚至不出现在发明家的思想中.除了他抛弃的一些组合——尽管它们也在某种程度上具有有用组合的特征——之外,在他有意识做出的组合的范围内,仿佛实际上从来也没有无用的东西.这一切就好像发明家是一位复试主考人,他只询问已经通过初试的候选人.

不过,我迄今所说的都是在读几何学家的著作时可以观察到或可以猜测到的东西,当然要动脑筋读.

现在,比较深刻地洞察和观察一下在数学家的心灵中发生了些什么.关于这一点,回忆一下我自己的情况是最好不过的了.不过,我将只限于讲述我是如何写出我的第一篇关于富克斯函数的论文的.我请求读者原谅,我正准备使用一些专门的表述,但是读者不要害怕它们,因为没有人强迫你理解它们.例如,我将说,我是在这样的情况下找到了这样一个定理的证明的.这个定理将有一个规范的名字,许多人都不知道它,不过这并不重要;心理学家感兴趣的不是这个定理而是发现这个定理的情况.

我曾用了两周时间力图证明不可能存在任何类似于我后来称之为富克斯函数的函数.我当时一无所知,我每天独自一人坐在我的办公桌前,呆一两个小时,尝试了最大的组合,什么结果也没有得到.一天夜晚,我违反了我的习惯,饮用了黑咖啡,久久不能入睡.各种想法纷至沓来;我感到它们相互冲突,直到成对地结合起来,也就是说,造成了稳定的组合.到第二天早晨,我已确立了一类富克斯函数的存在,它们来源于超几何级数,我只能写出结果,仅花费了几个小时.

接着,我想用两个级数之商把这些函数表示出来;这种想法完全是有意识的和深思熟虑的,与椭圆函数的类比指导着我.我问我自己,如果这些级数存在,它们必须具有什么性质,我毫不费力地获得了成功,形成了我所谓的 θ 富克斯级数.

恰恰在这时,我离开了我当时居住的卡昂,参加了矿业学校主办的地质考察旅行.沿途的景致使我忘却了我的数学工作.到达库唐塞后,我登上公共汽车去某个地方.当我的脚踩上踏板的一刹那,一种想法涌上我的心头,即我通常定义富克斯函数的变换等价于非欧几何学的变换,在我先前的思想中,似乎没有





什么东西为它铺平道路.我没有证明这一想法;我坐在公共汽车的座位上,继续进行已经开始的谈话,但是我感到它是完全可靠的.回到卡昂,为了问心无愧起见,我抽空证实了这一结果.

然后,我把注意力转向一些算术问题的研究,表面看来没有取得许多成果,也没有想到它们与我以前的研究有什么关系.我为我的失败而扫兴,于是前往海滨消磨几天时间,想一些其他事情.一天早晨,当我正在悬崖边散步时,一种想法浮现在我的心头,即不定三元二次型的算术变换等价于非欧几何学的变换,它正好具有同样的简洁、突然和直接可靠的特征.

返回到卡昂以后,我深思了这个结果,推导出一些结论.二次型的例子向我表明,存在着富克斯群,这些群不同于与超几何级数对应的群,我看到,我可以把 θ 富克斯级数理论应用于这些群,从而存在着一些富克斯函数,它们不同于当时我知道的、从超几何级数得到的函数.我自然而然地让我自己构造这一个函数.我向它们发起了系统的攻击,一个接一个地攻克了所有的外围工事.有一处外围工事无论怎样进攻还是岿然不动,只有攻陷它才能占领整个阵地.但是,我的全部努力乍看起来只是使困难清楚地呈现在我的面前,事情实际上就是这样.所有这些工作完全是有意识的.

紧接着,我要去瓦莱里昂山服军役,这样,我便从事截然不同的工作.一天,我正在大街行走,曾经使我感到困难的答案突然浮现在我的眼前.我无法立即深入探讨它,只是在服役结束后,我才开始继续研究这个问题.我已有全部元素,只需排列和整理它们.就这样,我一举写出了我最后的论文,丝毫没有感到有什么困难.

我只限于举这一个例子;多举也无用.关于我的其他研究工作,我可以讲出类似的情况,其他数学家在《数学教学》杂志中所给出的意见只会证实它们.

起初,最为引人注目的是顿悟的显现,这是先前长期无意识工作的明显征兆.在数学发明中,这种无意识工作的作用在我看来似乎是不可否认的,在其他不大明显的情况下,也可以发现它的痕迹.当人们研究一个艰难的问题时,在第一次进攻中往往达不到良好的效果.于是,人们或长或短地休息一下,坐下来重新工作.在起初半小时内,像以前一样,什么也找不到,然后一个决定性的想法突然浮现在脑海.可以说,有意识的工作之所以富有成效,是因为它被打断了,休息使思想生气勃勃、精力饱满.但是,很可能是这样:这种休息充满了无意识的工作,这种工作的结果后来向几何学家揭示出与我所举的例子恰好一样的东西.只是这种揭示不是在散步或旅行时发生,而是在有意识的工作期间发生的,不过它与有意识的工作无关;有意识的工作至多起了兴奋剂的作用,犹如它是一种刺激物一样,它激发了在休息时已经达到的结果,但依然是无意识的,尽管它采取了有意识的形式.

关于这种无意识工作的条件,还可以说明如下:如果一方面有意识的工作期间在它之前,另一方面有意识的工作期间又尾随其后,那么它就是可能的,而且肯定是富有成果的.这些突如其来的灵感(已经引用的例子充分证明了这一点)只有在有意识努力工作了若干天之后才会出现,尽管这些努力好像毫无成果,从中也没有得出什么好东西,而且所采取的路线似乎也是完全错误的.因此,这些努力并不像人们设想的那样一点结果也没有;它们驱动着无意识的机器,没有它们,无意识的机器就不会运转,也不会产生出任何东西.

对于在灵感之后需要第二个时期的有意识的工作,这还比较容易理解.必须使灵感的结果成形,从它们之中推导出直接的结论,整理它们,用语言表达出证明,而尤其必须验证这一证明.我已经谈过伴随灵感的绝对确实性的感觉;在所举的例子中,这种感觉不是骗子,而且,通常它的确不会骗人.但是,不能认为这个法则没有例外;这种感觉有时显得很逼真,也往往会欺骗我们,只有当我们企图进行证明时,我们才会发现这一点.当我早晨或晚上躺在床上处于半睡眠状态时,常有一些想法浮想联翩,我特别注意这一事实.

实际情况确是这样,现在,对于这种观念可作下述考察.无意识的自我,或者如我们所说的下意识的自我,在数学创造中起着重要的作用;这可由我们所说的情况中得出.但是,通常认为下意识的自我是纯粹自动的.现在,我们已看到,数学工作不是简单的机械工作,它不能用机器去完成,无论如何不能用机器圆满地完成它.这不只是应用法则的问题和按照某一固定的规律做出许多可能的组合的问题.这样得到的组合为数极多,但却是无用的、麻烦的.发明者的真正工作就在于在这些组合中进行选择,以便消除无用的组合,或者更确切地讲,躲避使他们陷入的困境,而且要使指导这种选择的法则极其精巧、极其微妙.要精确地陈述它们几乎是不可能的;与其说它们可以被表述出来,倒不如说它们可以被感觉到.在这些条件下,怎么能够设想一个机械地应用它们的筛子呢?

第一个假设现在呈现出来:下意识的自我决不劣于有意识的自我;它不是纯粹自动的;它能够识别;它机智、敏锐;它知道如何选择、如何凭直觉推测.我说什么呢?它比有意识的自我更清楚地知道如何凭直觉推测,因为它在有意识的自我失败了的地方获得成功.一句话,下意识的自我难道不比有意识的自我优越吗?你认识这个问题的充分重要性吧!布特鲁在最近的讲演中表明,它如何出现在十分不同的场合中,对它作肯定的回答会得出什么结论.

这种肯定的回答是因我刚才给出的事实迫使我们做出的吗?就我自己来说,我承认我不乐于接受它.那么,让我们重新审查一下那些事实,看看它们是否与另外的解释一致.

可以肯定,在无意识的工作延续了一段时间之后,以一种顿悟的形式呈现在我们脑海中的组合,一般说来是有用的、多产的组合,这似乎是初次印象的结





果.由此是否可以得出,通过微妙的直觉推测这些组合是有用的下意识的自我只不过是形成了这些组合呢?或者更确切地讲,下意识的自我是否形成了许多令人毫无兴趣的、依然是无意识的其他组合呢?

按照第二种方式观察下意识的自我,所有的组合都是由于下意识的自我的自动作用而形成的,不过只是有趣的组合才能闯入意识领域.这还是很神秘的.在我们无意识活动的无数产物中,只有一些被召唤通过阈限,其他的依旧在阈限之下,其原因何在呢?这种特权是简单的偶然性给予的吗?显然不是,例如,在我们感觉的所有刺激物中,只有最强的才能引起我们的注意,除非由于其他原因把我们的注意力吸引到不甚强的刺激物.更一般地讲,有特权的无意识的现象,即容许变成有意识的现象,就是直接或间接地最深刻地影响我们情感的现象.

关于教学证明,它似乎只能使理智感兴趣,当我们看到它也乞灵于情感时,可能会感到奇怪.这也许是忘记了数学的美感、数和形的和谐感、几何学的雅致感.这是一切真正的数学家都知道的真实的审美感,它的确属于情感.

那么,被我们赋予美和雅致这一特征的、能在我们思想上激起一种审美情感的数学实体是什么呢?它们是这样的实体:它们的元素和谐地配置,以致精神能够毫不费力地包容它们的整体,同时又能认清细节.这种和谐同时是我们审美需要的满足以及支持、指导我们思想的助手.而且,一个次序井然的整体摆在我们的双目之下,促使我们预见数学定律.于是,我们便达到下述结论:有用的组合恰恰是最美的组合,我意指最能使这种特殊情感着迷的组合,所有的数学家都知道这种情感,但是,对它的亵渎无知到常常嗤笑它的地步.

此外还发生了什么呢?在由下意识的自我盲目形成的大量组合中,几乎所有的都毫无兴趣、毫无用处;可是正因为如此,它们对审美感也没有什么影响.意识永远不会知道它们;只有某些组合是和谐的,从而同时也是有用的和美的.它们能够触动我刚才所说的几何学家的这种特殊情感,这种情感一旦被唤起,便会把我们的注意力引向它们,从而为它们提供变为有意识的机会.这只是一个假设,可是在这里有可以证实它的观察材料:当一个顿悟抓住数学家的思想时,它往往不会欺骗他,但是正如我说过的,它有时也不经受证据的检验,好了,我们几乎总是注意到,这种人为的观念倘若是真实的,它就会使我们对于数学雅致的自然情感得到满足.

因此,正是这种特殊的审美感,起着我已经说过的微妙的筛选作用,这充分地说明,缺乏这种审美感的人永远不会成为真正的创造者.

可是,困难并没有完全消失.有意识的自我严格地受到限制,至于下意识的自我,我们不知道它的限制,这就是为什么我们乐于假定,它在短时间内能够做出的各种组合比有意识的自有整个一生能够完成的组合还要多.可是,这种限

制是存在的.下意识的自我能够形成所有可能的组合,其数目之多远非想象力所能及,这种说法可信吗?不管怎样,这似乎是必要的,因为假使它仅仅产生一小部分组合,假使它随意地做出它们,那么我们在它们中能够选择、可以发现的有效的组合的机遇就会很少.

有意识的工作总是在所有富有成果的无意识的劳动之前,也许我们应当在有意识的工作的初期寻求解释.请允许我粗略地作一比较,把我们组合中的未来元素想象为伊壁鸠鲁(Epicurus)的带钩原子吧.在精神完全休眠时,这些原子是不动的,也可以说,它们钩住了墙壁;因此,这种完全的休眠可以无限地延续下去,没有相遇的原子,从而在它们之间也没有任何组合.

与此相反,在表面的休眠和无意识的工作期间,它们中的某些原子脱离墙壁并开始运动.它们通过封闭它们的空间(我正要说房间)向各个方向传播,犹如一群蚊虫,或者你如果喜欢学术上比喻的话,它们就像气体运动论中的气体分子.于是,它们的相互碰撞可以产生新的组合.

初期的有意识的工作有何作用呢?显而易见,它使这些原子中的某一些可以运动,它把它们从墙壁上卸下来并使它们自由活动.我们认为,我们之所以没有做出有效的组合,是因为要把这些原子集合起来,就要使它们以无数不同的方式运动,即便如此还是找不到满意的集合.但是,通过我们的自由意志使这些原子开始运动之后,它们就不会返回到它们的初始状态.它们自由地继续它们的运动.

好了,我们的自由意志并非随意地选择它们,它追求一个完全确定的目的.因此,可动的原子并非无论什么样的原子;它们是我们可以合理地期望从中得到所要求的答案的原子.于是,可动原子经受碰撞,从而使它们进入它们之间的组合,或者与在它们的进程中撞击到的其他静止的原子形成组合.我再次请求原谅,我的比喻是很粗糙的,但是我不知道如何用其他方法使我的思想得以理解.

不管情况如何,有形成机遇的组合仅仅是这样一些组合,即其中的元素至少有一个是由我们的自由意志自由地选择出的那些原子之一.现在,很明显,在这些组合中,可以找到我所谓的有效的组合.也许这是一种减少原来假设中的悖论的途径.

还有另外的意见.我们仅应用固定的法则,全部做出了有点冗长的运算,无意识的工作永远也不会为我们提供这个结果.我们可能以为,完全自动的下意识的自我特别适于这种工作,这是一种全部机械的方法.也许我们夜晚思考乘法因子,我们醒来时必定希望找到预先做出的乘积,或者还希望代数运算——





例如证明——可以无意识地被做出来.正如观察所证明的,根本没有这种事.从这些灵感中,从无意识的工作的成果中,人们可望得到的一切只是这样的运算的出发点.至于运算本身,必须在紧随灵感之后的有意识的工作的第二个期间完成,人们在此期间验证这一灵感的结果,推出它们的结论.这些运算的法则是严格的和复杂的.它们要求纪律、注意力、意志,因而要求意识.相反地,在下意识的自我中,则是由我所说的自由统治着,倘若我用自由这个名称称呼绝对缺乏纪律和源于机遇的混乱的话.不过,这种混乱本身却容许未曾料到的组合.

我将作最后的评论:当我在上面发表某些个人的观察材料时,我谈到我在不由自主地工作时的令人激动的夜晚.这样的情况是经常发生的,不过大脑的反常活动不必要由我曾提到的物质刺激物引起.在这样的情况中,人们在他无意识的工作中呈现出的东西似乎可以部分地被过分激动的意识所领悟,可是这并不改变无意识的工作的本性.于是,我们不甚明确地理解了两种机制——如果你愿意的话,也可以说是两种自我的工作方法——的区别是什么.而且,在我看来,我从中能够做出的心理学观察材料似乎在它们的总轮廓上证实了我刚才提出的观点.

的确,这些观点需要证实,因为不管怎样,它们是而且确实依然是真正的假设;对这些问题的兴趣如此之大,以至于我一点也不后悔向读者提出了上述观点.

4 数学和逻辑[①]

几年前,我有机会提出了某些关于无限的逻辑的观念,谈论了无限在数学中的作用和自康托以来由它所构成的应用.我解释过,我为什么不认为某些推理方法是合理的,而许多著名的数学家却相信它们可以使用.不用说,我招来了一些尖锐的答辩.这些数学家不相信他们错了;他们坚信他们有权做他们曾经做过的事情.讨论拖了下去,这并不是因为不断地提出了新的论据,而是因为我们继续在同一个圈子里团团转,每个人都重复着他刚刚说过的话,似乎没有听到对手已经说过什么.在每一个场合,我都要就所争论的原理提出新的证据,可以说是为了不致遭到大家反对;但是,这种证据总是相同的,几乎未加修改.因此没有得出结论.假如我说我感到意外的话,那是传达了假象,其实我的心理是亮堂的.

在这些条件下,再次重复同样的论据似乎是不可取的,我可以给这些论据

① 注:原载[法]庞加莱著,最后的沉思.李醒民译,范岱年校.北京:商务印书馆,1995,75~86.

以新的形式,但却基本上不会改变它们,因为在我看来好像是我的对手甚至没有试图去拒绝它们.寻求造成这种截然不同的观点的智力差别的起源似乎是可取的.我刚刚说过,这些不能缩小的分歧并不使我感到惊讶,我从一开始就已经预见到分歧.但是,这并未免除我们寻求解释;在反复经验之后,预见事实是可能的,还被紧紧催逼着要去解释它.

因此,让我们尝试从纯粹客观的观点来研究一下两个对立学派的心理学,就好像我们自己不是这两个学派的成员,就好像我们正在讲述两窝蚂蚁打仗一样.首先,我们将看到,数学家在他们考虑无限的方式方面存在着两种对立的倾向.在一些数学家看来,无限是由有限导出的;无限之所以存在,是因为有无限多可能的有限事物.对另一些数学家来说,无限在有限之前就存在着;有限是从无限切下一小段而得到的.

一个定理必须能够证明,但是由于我们自己是有限的,我们只能够处理有限的对象.这样一来,即使无限的概念在定理的陈述中起作用,但是在证明中必须不涉及它;否则,这种证明将是不可能的.我将引用下面的定理作为一个例子:素数集无界;级数 $\sum \frac{1}{n^2}$ 是收敛的等.这些例子中的每一个都能够化为只包含有限数的等式和不等式.这些定理带有无限的特征,并不是因为一种可能的证明本身带有无限的特征,而是因为可能的证明在数目上是无限的.

在陈述定理时,我断言它的所有证明将为真.这被理解为,并非所有的证明全部给出了.还有一些我认为是可能的证明,因为它们大概只需要有限长的时间,但是它们实际上是不可能的,由于它们可能需要多年的工作.我相信,要是我们能够设想一些富有而愚蠢的人(他们足以雇用充分多的帮手)企图完成它,那就好了.但是,作为定理证明的真正目的,它又使这种蠢事变得没有必要.

不能得出任何可验证结论的定理有意义吗?或者,更普遍地讲,任何定理除了与它有关的证明外还有意义吗?这正是数学家有分歧的地方.第一个学派的那些数学家说没有,我将称他们为实用主义者(因为有必要给他们取一个名字);当一个定理在没有给他们以验证它的方法的情况下而引起他们的注意时,他们在其中看到的只是不可理解的冗词赘句.他们愿意考虑的只是能够用有限数目的词定义的对象.在一个论据中,当提到作为满足某些条件的对象 A 时,他们理解满足这些条件的对象,不管用来完成它的定义的词汇可能是什么,尽管这些词在数目上是有限的.

另一个学派的数学家不想承认这一点,我将把他们简称为康托主义者.一个人不管他多么健谈,他在他的一生中也不能说十亿以上的词汇.因此,我们将从科学中排除其定义包含十亿零一个词的对象吗?如果我们不排除它们,我们





为什么要排除那些只能够用无限数目的词定义的对象呢,这是由于第一类定义的表述像第二类定义的表述一样超越了人类的范围吗?

不难理解,这个论据使实用主义者大为扫兴;不管一个人多么健谈,人类还将更为健谈,因为我们不知道人类将延续多么长的时间,我们不能预先限制人类的研究范围.我们仅仅知道,这个范围将总是有限的;即使我们也许能够确定人类消亡的日期,但是还有其他天体上的智慧生物,能够继续从事在地球上留下的未完成的工作.而且,实用主义者在设想比我们更健谈,而且还保留着某些人性的人类时,他们也许并不疑虑不安;他们不愿就关于在有限长的时间内能够思考无限多词汇的一些无限健谈的神灵的假说进行争论.另一方面,其他人认为,客体与能够谈论或思考它们的任何人类或任何神灵无关地大量存在着;我们能够在这种贮存中自由地选择:我们无疑没有足够的欲望或充裕的金钱来购买每一样东西;但是库存货物却与买主的资财毫不相干.在细节上的所有各种分歧就起因于这种最初的误解.

让我们以策默罗定理为例,按照该定理,空间能够变换为良序集.康托主义者将被证明的严格、真实或明显所迷住,实用主义者将回答:

"你说你能把空间变换为良序集.好吧,变换它!"

"那需要花费很长时间."

"那么,你至少要向我们证明,某个有足够的时间和耐性的人能完成这种变换."

"不,我不能证明,因为实行变换的操作数目是无限的;它甚至比阿列夫零(Aleph-zero)还要多."

"你能够指出容许空间是良序的定律如何用有限数目的词来描述吗?"

"不能."

于是实用主义者得出结论:该定理或者没有意义,或者为假,或者至少未被证明.

实用主义者采用外延的观点,康托主义者采用内涵的观点.当涉及一个有限的集合时,这种区分只有对形式逻辑理论家来说才是有意义的;但是,当涉及无限的集合时,这种区分对我们来说似乎具有更深远的意义.如果我们采用外延的观点,那么集合可以通过新数的相继添加而形成;我们能够把旧对象结合起来构造新对象,然后用这些新对象构造更新的对象;如果集合是无限的,正是因为不存在停下来的理由.

另一方面,从内涵的观点来看,我们从其中具有预先存在的对象的集合开始,这些对象乍看起来似乎是没有区别的,但是我们最终能分辨出它们中的几

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

个,因为我们标记了它们,并且把它们排列在抽屉里.但是,对象在标记前就存在着,集合也会存在,即使也许没有把它们进行分类的管理员.

对于康托主义者来说,基数的概念没有包含任何秘密.当两个集合能够排列在相同的抽屉时,它们就具有相同的基数;事情不会更容易了,由于两个集合预先存在着,同样可以认为与负有排列对象任务的管理员无关的抽屉内的集合预先存在着.对于实用主义者来说,情况并非如此.集合没有预先存在;它每天都增长着;新对象不断地变得与它有关,如果不涉及预先已经分类的对象概念和它们分类的方式,人们也就不能定义这个集合.每逢一个新的获得物时,管理员都可能被迫打乱抽屉,以便找到一种按适当顺序配置它的方法;两个集合是否能够排列在相同的抽屉内,这将永远不会为人所知,因为总是要担心,打乱它们将是必要的.

例如,实用主义只承认能够用有限数目的词定义的对象;能够用语句描述的可能的定义总是能够用从一到无限的寻常数来计数.根据这种推断,也许只存在可能的单重无限基数,即阿列夫零数.可是,我们为什么说连续统的幂不是整数幂呢?是的,给出我们能够用有限数目的词定义的空间中所有点后,我们就能够想象一个定律,该定律本身能够用有限数目的词来描述,而且能在这些词和整数集之间建立起对应.但是,现在让我们考虑其中包含着这个对应定律概念的语句.不久前,这些语句没有意义,因为这个定律还没有被发明出来,它们不能用来定义空间的点.现在,它们已获得了意义;它们将容许我们定义空间的新点.但是,这些新点将在已经采纳的分类中找不到任何位置,这将迫使我们打乱它.在实用主义看来,当我们说连续统的幂不是整数幂时,我们的意思就是这样.我们意味着,在这两个集之间不可能建立摆脱这类混乱的对应定律;而在涉及直线和平面的例子中,就有可能做到这一点.

其次,实用主义者没有肯定,是否无论什么集恰当地讲都具有基数;或者,给定两个集,是否总有可能知道,它们是否具有相同的幂,或者一个幂是否比另一个幂大.从而他们被导致怀疑阿列夫的存在.

分歧的另一个来源起因于构想定义的方式.存在着各类定义;存在着通过近缘的类和不同的种,或者通过合成能够导出的直接定义.

让我们附带注意一下,在不能定义特殊的事物,而只能定义整个种的意义上,存在着不完全的定义.它们是合理的,它们甚至是最为频繁使用的定义.但是在实用主义者看来,有必要在其中理解特殊对象的集,这些对象满足该定义,并且最终能够用有限数目的词来定义.因为康托主义者的这种限制是人为的,而且没有意义.





如果仅存在直接定义,那么纯粹逻辑的重要性就不可能引起争议.于是,无论在什么命题中,都可能用它的定义代替每一个术语.当完成这种代替时,要么该命题不能简化为等同,从而不能是纯粹逻辑证明,要么它能简化为等同,从而只不过是或多或少精巧伪装起来的同义反复.

但是,还有另外一类定义,即用公设来定义.一般地,我们总是知道,被定义的对象属于一个类;但是,当陈述特定的差别成问题时,那就不直接陈述,而借助于被定义的对象必须满足的"公设"来陈述.就这样,数学家能够借助于显方程 $x=f(y)$ 或隐方程 $F(x,y)=0$ 来定义量 x.

只有当所定义的对象的存在被证明时,用公设定义才有价值.用数学语言来说,这意味着该公设没有隐含矛盾;我们没有权利忽略这个条件.要么必须承认,由于一种信念的作用,无矛盾是直观真理、是公理——可是这样就必须认清我们正在做的事情,铭记我们已经扩大了不可证明的公理的一览表——要不然就必须借助于法则或公设或利用递归推理来构造形式证明.尽管当涉及直接定义时这种证明并非不大必要,但是它一般来说却比较容易.

一些实用主义者可能更为严谨;为了使他们认为定义是合理的,在术语上不导致矛盾是不充分的;按照我在上面试图定义的他们的特殊观点,他们要求定义要有意义.

不管事情可能怎样,在通过公设引入定义后,逻辑将依然是无结果的吗?在给定一个命题后,我们不再能够在其中用定义代替一个术语.我们能够做的一切就是在命题和作为它的定义的公设之间消除这个术语.如果这种操作是按照所谓的逻辑消去法则进行的,那么它就不会导致等同,因为该命题不能借助于纯粹逻辑来证明.如果它导致等同,那正是因为该命题只不过是同义反复而已.我们不需要在我们不久前所作的结论中改变任何东两.

但是,还有第三类定义,这是实用主义者和康托主义者之间新误解的起源.这些定义也是通过公设来定义,但是公设在这里是被定义的对象和一个类的所有个别对象之间的关系,被定义的对象本身被假定是这个类中的一个元(或者人们假定它们本身只能够用要被定义的对象来定义的那些对象是这个类的元).如果我们假定下述两个公设,所发生的情况就是这样.

X(被定义的对象)以这样的方式与类 G 的所有元有关.

X 是 G 的元.

要不然,假定下述三个公设:

X 以这样的方式与类 G 的所有元有关.

γ 以这样的方式与 X 有关.

γ 是 G 的元.

在实用主义者看来,这个定义隐含着循环论证.在不知道类 G 所有元的情况下,从而在不知道这些元之一 X 的情况下,就不可能定义 X.康托主义者不承认这一点;类 G 被给定,从而我们知道它的所有元.作为目的,该定义仅仅必须从这些元中区分出一个元,它与它的所有同伙元具有所描述的关系.

"不",他们的反对者回答说:"类的知识不会导致你认识它的所有元;它只不过向你提供了构造所有元的可能性,或者更确切地讲,提供了构造你所希望的那么多的元的可能性.它们将只有在它们被构造之后才存在;也就是说,在它们被定义之后才存在;X 只有借助于它的定义才存在,只有 G 的所有元,尤其是 X 预先已知,它才具有意义."他们附加道:"说下面的这些话可能是无用的;例如说什么用它对于 X 的关系来定义 X 并不是循环论证;说什么总之这个关系是能够被用来定义 X 的公设;因为必须预先确定,这个公设不隐含矛盾.但是,那不是通常在这种类型的定义中所要做的事情.我们首先证明,无论类 G 可以是什么,假定所有它的元都已知,它也许由于这个类而具有所述的关系:也就是说,这个对象的存在并不导致矛盾.在这里,可能留下来的是要证明,在这个对象的存在和假说之间没有矛盾,这个对象本身是该类的元."

争论可能会继续一个很长的时间;但是,我乐于强调的观点是,如果容许这类定义,那么逻辑就不再是无结果的了,而且证明就是用预定证明命题的方式来系统表述大量论据,这些命题决不是同义反复,因为有些人仍拿不准它们是否错了.因此,我们为一个词所能具有的能力而惊奇.在这里,有这样一个对象,在它被命名之前,从它之中连什么东西也不能推导出来;它所需要的一切就是取个名字,这名字创造了奇迹.这如何能够发生呢?因为给它取个名字,我们就已隐含地断言,该对象确实存在着(也就是说,摆脱了所有矛盾),它完全被确定了.但是,在实用主义者看来,我们根本不知道这一点.事实上,使这个证明变得毫无结果的机制是什么呢?那是很简单的;我们假定,被证明的命题为假,我们证明这导致与对象 X 存在的事实相矛盾.只要我们肯定它的存在,而且只要我们知道该对象完全被确定了,这就是合理的.实际上,要是 X 是通过定义从类 G 推出就行了;其次,要是类 G 是通过包括对象 X 和能够从类 G 中推导出的所有其他元在内而变得完全就行了;如果这样而变得完全的类称为 G',如果我们把能够通过定义,并且用与 X 从 G 推导出来的同一方式而从 G' 推导出的元称为 X',那么就必须确信 X' 等同于 X.如果情况并非如此,如果通过假定被证明的命题为假,我们便被引导到两个矛盾的陈述

$$\varphi_1(X)=0, \varphi_2(X)=0$$





那么,我们怎样才会知道,在两个陈述中所涉及的是同一个 X 呢?如果 X 包含在一个陈述中,而 X' 包含在另一个陈述中,那么两个命题就可写成

$$\varphi_1(X)=0,\varphi_2(X')=0$$

一般说来,它们不再是矛盾的.

为什么实用主义者因此会提出这种异议呢?因为对于他们来说,类 G 似乎只是能够无限增加的集合,无论何时新的元都能形成,它们具有适当的特征.于是,G 从来也不能像康托主义者所做的那样不可改变地被安排,从而我们无法肯定,借助于新的附加物它将不变为 G'.

我力求尽可能清楚、尽可能公正地解释两个学派数学家的分歧的本质.对我来说,这似乎是我们已经能够领悟出的真正的原因.两个学派的科学家具有对立的思想倾向.我称之为实用主义的那些人是观众论者,而康托主义者是实在论者.

存在着一种能够证实这种观点的东西.我们看到,正如我所说的,康托主义者(让我使用这个方便的术语吧,尽管我在这里不希望谈论步康托后尘的数学家,甚至也许不想谈论那些认为他们与康托一致的哲学家,而只想谈谈在独立的形式方面具有同一倾向的人)不断地谈到认识论,即科学的科学.这种认识论完全与心理学无关,这一点已被充分地理解;也就是说,它必须告诉我们,假使没有科学家的话,究竟科学是什么;我们必须研究科学,这当然没有假定不存在科学家,但至少是没有假定存在科学家.于是,不仅自然是独立于试图研究它的物理学家的实在,而且物理学本身也是一种实在,即使没有物理学家,它也存在着.事实上,这就是实在论.

实用主义者为什么不肯容许不能用有限数目的词来定义的对象呢?这是因为他们认为,对象只有在它能用心智构想时才存在.对象不能用独立于有能力思考的人的心智来构想.实际上,在这里存在着观念论.既然有理性的主体是人,或者是类似于人的某种生物,因而是有限的存在,所以无限除了有创造我们所希望的那么多的有限对象的可能性外,它没有别的意义.

这样,我们可以做出某种特殊的评论.实在论者通常采取物理学家的观点.他们断言物质对象、或个体灵魂、或他们所谓的实物的独立存在.在他们看来,世界在人创生之前就存在着,甚至在生物创生之前就存在着;即便没有上帝,或没有任何理性生物,世界还会存在.这是常识的观点,只有通过沉思我们才能抛弃它.物理实在论的支持者一般说来都是有限论者.至于谈到康德的二律背反问题,他们对该论题亦步亦趋;他们相信世界是有限的.例如,这是伊夫琳(Evellin)先生的观点.另一方面,观念论者并没有同样的顾忌,他们已充分准备好

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

赞同对立的观点.

可是,康托主义者是实在论者,甚至在涉及数学实体的地方也是如此.在他们看来,这些实体似乎具有独立的存在;几何学家并没有创造它们,他只是发现它们.因此,这些对象可以说在没有现存的情况下就存在着,因为它们能够归结为纯粹的本质.但是,由于这些对象就其本性而言在数目上是无限的,因此数学实在论的支持者与观念论者相比,他们是更大程度的无限论者.在他们看来,无限由于在发现它的心智之前就存在着,因而它不再是生成(Becoming).不管他们承认还是否认无限,他们必须因此而相信实无限.

我们在这里辨认出柏拉图(Plato)的理念论;看到把柏拉图归入实在论者之中可能似乎是奇怪的.不过,没有任何学说比柏拉图主义更强烈地与当代观念论相对抗了,尽管这种学说也远离物理实在论.

我从未见到有比埃尔米特(Hermite)更为实在论的数学家(在柏拉图的意义上的实在论),我还必须承认,我从来也没有遇见一个比他更反对康托主义的人.在这里,似乎存在着表面上的矛盾,之所以更加如此,是由于他乐意重复说:"我之所以是一个反康托主义者,因为我是实在论者."他因创造对象而不是满足于发现它们而责备康托.毋庸置疑,由于他的宗教信念,他认为,希望毫无困难地深入到只有上帝才能够理解的领域,而不等待上帝向我们一个接一个地揭示它的秘密,这是大逆不道的行为.他把数学科学和自然科学加以比较.在他看来,博物学家企图猜测上帝的秘密,而不通过经验来了解,这对神圣的上帝不仅是放肆的,而且是无礼的.在他看来,康托主义者似乎想要以同样的方式在数学中行动.这就是为什么他在实践上是观念论者,而在理论上是实在论者.存在着一个已知的实在,它在我们的外部,不依赖于我们;但是,我们关于它所能知道的一切都依赖于我们,于是这一切只不过是生成,是一种相继获得的层次.其余的东西是实在的,却是永远不可知的.

无论如何,埃尔米特的情况是一个孤立的例子,我不希望进一步停留在它上面.不论何时,在哲学中总是存在着对立的倾向,这些倾向似乎并没有处于和解的边缘.毫无疑问,这是因为存在着不同的心灵,我们不能改变这些心灵中的任何东西.因此,没有希望看到在实用主义者和康托主义者之间建立起和谐.人们没有取得一致,因为他们讲的不是同一种语言,有的只是彼此都不能学会的语言.

然而,在数学中,人们通常可以彼此了解:但是,这恰恰是由于我已经称之为证明的东西.这些证明在没有上诉的情况下就宣布判决,在它们面前,整个世界都得屈从.但是,不管在什么地方,如果缺乏这些证明,数学家就一点也不比





头脑简单的哲学家高明.当必须了解一个定理在无法证明的情况下能否具有意义时,由于根据定义我们不允许我们自己去证明它,谁能够判断它能否有意义呢?除了因矛盾而使对手走投无路外,不会有其他办法.但是,人们已经尝试做了实验,却未获成功.

许多二律背反都被指出来了,不一致依然存在;没有一个人被说服.总有可能通过改变论据使自己摆脱矛盾;我指的是通过区别.

第三章

庞加莱的数学贡献①

庞加莱1854年4月29日生于法国南锡,1912年7月17日在巴黎逝世.他是一位著名数学家、数学物理学家和天体力学家.

19世纪数学的发展是从高斯的高大身影开始;而在另一个同样伟大的天才庞加莱的支配下宣告结束.二者都是非常了不起的全能的数学家,对天文学和数学物理学都做出了巨大的贡献.如果庞加莱在数论方面的发现不能同高斯相提并论的话,那么他在函数论上的成就至少是可以同高斯媲美的,就是说到椭圆函数和模函数理论也是这样,这些理论当然无疑归功于庞加莱,并且是他在这一领域内的重大发现,尽管他在生前没有将这些工作发表出来.如果说高斯是微分流形理论的创始者,那么庞加莱在建立代数拓扑理论方面起同样的作用.最后,庞加莱还是微分方程理论方面最重要的人物,并且是继牛顿之后在天体力学方面做了最出色的工作的数学家.高斯和庞加莱两人都只有极少的学生,他们喜欢独自工作,但他们之间的相似之处仅此而已;高斯非常不愿意发表他的发现,而庞加莱所发表的论文的数目接近500篇,还不包括许多书和他在巴黎大学讲学时发表的讲演.

① 本章作者:让·迪多内.译自 Dictionary of Scientific Biography XI 中 Jules Henri Poincaré.井竹君,译.章祥孙,校.



庞加莱的双亲都出生于中上阶层家庭,这两家几代都居住在法国东部的洛林.他的祖父有两个儿子:庞加莱的父亲 Léon 是一位医生,并且是南锡大学医学教授;另一个儿子 Antoine 曾在高等工科学校学习过,以后在工程兵部队中居要职.Antoine 的一个儿子 Raymond 曾几度任首相,在第一次世界大战期间任法兰西共和国总统,另一个儿子 Lucien 在大学里担任高级行政职务.当庞加莱还只是中学生时,他的数学才能就显露出来了,他曾荣获全法国公立中学生数学竞赛一等奖,并以班中的最优成绩于 1873 年进入高等工科学校;据说他在南锡就学时,他的教授称他为"数学巨人".毕业后,他在矿业学院继续修工程方面的课程,同时撰写数学博士论文,于 1879 年获得了博士学位.之后不久,他开始在 Caen 大学任教,1881 年成为巴黎大学的教授,以后就一直在那里教书,直到 1912 年过早地去世.当他 33 岁时,以那样年轻的年龄被选进科学院,又于 1908 年选进法兰西学院.他曾无数次地获得国内和国外颁发的奖金和荣誉.

1 函数论

还在 30 岁以前,庞加莱就以他的单个复变量的"自守函数"的划时代发现而闻名于世(他称这些函数为 fuchsian 和 kleinean 函数),模函数和超几何方程的解的研究曾给出了定义在复平面上一个开连通子集 D 上的解析函数的例子,它们在 D 到其自身的变换群 G 下是不变的,这变换形如

$$T: z \to (az+b)/(cz+d) \quad a,b,c,d \text{ 为常数}, ad-bc \neq 0 \qquad ①$$

G 是"真不连续",也就是说,对于一串变换,$T_n \in G$,D 中不存在一点 z,是 D 中某点 z' 在 T_n 下的象的无穷序列(不等于 z)的极限,例如,变换①组成一模群,这里 a,b,c,d 为整数,且 $ad-bc=1$;D 为上半平面 $\phi z>0$,可由基本区域 $|z|\geqslant 1, |Rz|\leqslant \frac{1}{2}$ 的所有变换无重叠地覆盖.庞加莱巧妙地利用非欧几何,指出,对于任一形如①的真不连续变换群 G,存在类似的基本区域,此基本区域以直线或圆的一部分作为边界,并且它在 G 的元素的变换下无重叠地覆盖 D.反之,角和边满足某些明确条件的任何一个"圆多边形",一定是形如①的某个真不连续变换群的基本区域.开集 D 可以是半平面 $\phi z>0$,或是圆的外部或内部;当不取这种形式时,它的边界可以是一完备的无处稠密集,或是一曲线,在其任一点上都没有切线或曲率.

其次,仿照椭圆函数论中的魏尔斯特拉斯级数,庞加莱指出,对于一给定的群 G,以及一个在 D 的边界上无极点的有理函数 H,定义级数

$$\Theta(z) = \sum_k H(T_k \cdot z)(c_k z + d_k)^{-2m} \qquad ②$$

这里变换

$$T_k: z \to (a_k z + b_k)/(c_k z + d_k)$$

是变换 G 的列举，m 为一足够大的整数，则除了在变换 G 下成为 H 的极点的点外，级数②是收敛的；由此定义在 D 上的亚纯函数 Θ 显然对群 G 的任一变换①满足关系

$$\Theta(T \cdot z) = (cz + d)^{2m} \Theta(z)$$

对应于同一整数值 m 的两个这样函数的商，庞加莱称之为 Thetafuchsian 或 Thetakleinean，这商给出一个自守函数（D 上的亚纯函数）. 容易证明，任意两个自守函数 X, Y（在 D 上为亚纯的且相应于同一群 G）满足一个"代数"关系 $P(X, Y) = 0$，这里曲线 $P(x, y) = 0$ 的亏格等于齐次空间 D/G 的拓扑亏格，且能从（正如庞加莱所指出的）G 的基本区域明确地计算出来. 而且，如果 $V_1 = (\mathrm{d}X/\mathrm{d}z)^{\frac{1}{2}}$，$V_2 = 2V_1$，$V_1$ 和 V_2 为二阶线性微分方程

$$\frac{\mathrm{d}^2 V}{\mathrm{d}X^2} = \varphi(X, Y) V$$

的解，这里 φ 是 X, Y 的有理函数，则自守函数 X 可由关系 $z = V_1(X)/V_2(X)$ "求逆"而得到. 这一性质是庞加莱研究的出发点，而这些研究是追随了 I. L. Fuchs一篇研究二阶方程 $y'' + P(x)y' + Q(x)y = 0$ 的论文，Fuchs 的方程中 P, Q 为有理系数，这方程两个解的商的逆给出一亚纯函数；庞加莱就是由此来给他的自守函数选名字的.

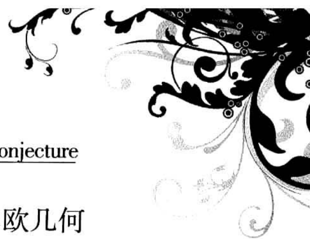
但是庞加莱没有就此止步. 他注意到在构造他的 fuchsian 函数时引入许多容许连续变化的参数，他猜想，适当地选取这些参数，对于"任意"的代数曲线 $P(X, Y) = 0$ 可得到一借助于 fuchsian 函数的参数表达式，另外，对任意阶的齐次线性方程

$$y^{(n)} + P_1(x) y^{(n-1)} + \cdots + P_n(x) y = 0$$

其中，P_j 为 x 的代数函数，也可能用 zethafuchsian 函数来表达其解（这样一个函数 F 在空间 C^p 中取值；换言之，这是 P 个纯亚纯函数的系统，并使得对它所相应的 fuchsian 群 G 中的任意变换，有 $F(T, z) = \rho(T) \cdot F(z)$，这里 ρ 是由 G 到 C^p 的一个线性表示），他用来探索证明这些结果的"连续性方法"在当时是不可能太严密的，因为在 19 世纪 80 年代这样早的时候还缺乏适当的拓扑概念和结果；但是有了 Brouwer 的拓扑学基本定理之后，利用稍微不同的方法就能给出正确的证明.

许多文章写到 C·F·克莱因（C. F. Klein）和庞加莱在发现自守函数问题上的"竞争". 实际上从来没有过真正的竞争，从一开始，克莱因就同他差一段距离. 可以肯定，在 1879 年，克莱因是知道了以前所有的关于特殊自守函数的文章，在这方面，他也通过数篇有关椭圆函数变换的漂亮论文而做出了自己的贡献. 尤其是他不可能不注意到这些函数的基本区域和非欧几何之间的联系，因为正





是他继 Cayley 和 Beltrami 之后澄清了以“庞加莱半平面”为特例的各种非欧几何的欧氏“模型”的概念.

另一方面,难以置信的是,当庞加莱开始其研究时,他对数学文献却都不熟悉.除了 Hermite 在模函数方面的工作之外,他几乎不知道任何其他东西;他肯定不曾读过黎曼的书,据他自己说,他甚至不曾听到过“狄利克雷原理”,而他只是在数年之后以一种推测的方法使用过这一原理.尽管如此,似乎在克莱因的思想中,没有产生过庞加莱把一个基本区域同任意 fuchsian 群相联系的概念,关于“利用”非欧几何的观念也是一样,因为直到 1880 年克莱因在他关于模函数的论文中从没有提及过这些思想.克莱因在他给庞加莱的信中提出的问题之一就是他是如何证明“θ 级数”的收敛性的.只是在领悟到庞加莱正在寻找一个所有的代数曲线的亚纯函数参数表达式的定理以后,克莱因才开始动手证明这一点,从而成功地构思出一个独立于庞加莱的证明.但他应用的方法是与庞加莱类似的(同样缺乏严格性).

单复变量的自守函数的一般理论是庞加莱没有留下什么工作可给他的后继者去做的少数数学分支之一.不能把单复变量的自守函数“自然”地推广到多复变量的情形.目前的认识提示我们,这理论可能同 E. Cartan 的对称空间 G/K 理论相关(G 为半单纯实 Lie 群,K 为 G 的最大紧致子群),以及同在 G/K 上运算,并使得 G/Γ 有有限测度的 G 与离散子群 Γ 有关(C. L. Siegel).而从这观点出发,作为庞加莱理论基础的群 $G = SL(2, R)$ 显得非常特殊,这是个唯一的单纯 Lie 群,在这里离散子群 Γ 的共轭类依赖于连续的参数(A. Weil 的刚性定理).因而不能再用庞加莱爱用的“连续性”方法了;事实上,已知的离散群 $P \subset G$ 是用算术方法来定义的,使得 G/Γ 有有限测度,这样多变量的自守函数比一个变量的自守函数更接近数论(在这方面,庞加莱很早就注意到由三元二次型的算术理论中导出的特殊“fuchsian 群”,以及相应的自守函数的特殊性质).

自守函数的理论仅是庞加莱对解析函数理论的许多贡献之一,这些贡献的每一个都是一些广博的理论的起点.他在 1883 年的一篇短文中首先研究了(用魏尔斯特拉斯方法分解成基本因子这个性质来定义的)整函数的亏格同整函数的 Taylor 展开式的系数或这函数绝对值的增长率之间的联系;和 Picard 定理合起来,并由阿达马和 E·波莱尔的结果,形成了整函数和亚纯函数的整套理论,这一理论经 80 年后还在发展.

自守函数给出了奇点,构成一完备的无处稠密集的解析函数的第一个例子,同时也给出了有奇点曲线的函数的例子.庞加莱借助于有理函数的级数 $\sum_n \frac{A_n}{z - b_n}$,给出了形成这类函数的另一种一般方法,E·波莱尔和 A·当儒瓦较迟发展的单演函数论就是从这一般方法得到的.

代数曲线的参数化定理也是由自守函数理论得到的一个结果,这一结果使庞加莱于1883年得到一般的"单值化"定理,这一定理等价于在平面或一开圆盘上的任意单连通非紧致的黎曼曲面的保角映象的存在性.这时,他明白了这一问题是狄利克雷问题的一个推广,庞加莱首次引入了利用紧致区域的递增序列来"穷举"黎曼曲面以及由极限过程来获得保角映象的思想.在当时要给出一个完全满意的证明再次遇到了困难,庞加莱自己和克贝不得不在1907年回到这问题上,直到这个问题被认为完全解决时为止.

克贝(Koebe, Paul, 1882—1945),德国数学家.

庞加莱甚至还是多复变量解析函数理论的倡导者——这一理论在他之前实际上还不存在.他的第一个工作是给出了两个复变量的亚纯函数 F 是两个整函数的商的定理,这是他在1883年巧妙地将狄利克雷原理应用到函数 $\log|F|$ 上来证明的;在较后的一篇论文中(1898),他深入研究了任意多个复变量的"复调和"函数,并应用到 Abelien 函数理论中.到了更晚些(1907),哈托格斯的定理发表后,他指出了由于两个复变量函数的"保角映象"概念的推广而引出的一些全新的问题.这些问题是我们今天所知的令人印象深刻的"解析几何"(或者是解析流形和解析空间理论)的胚芽,随后有 Causin, Hartog 和 E. E. Levi 在1914年前关于这方面的先驱工作;1930年代 H. Cartan, K. Oka, H. Behnke 和 P. Lelong 的工作;1945年后由上同调的思想给了此理论以极大的推动力.

哈托格斯(Hartogs, Friedrich M., 1874—1943),比利时数学家.

最后,庞加莱第一次对多复变函数的多重积分的"残数"概念做出一个满意的推广,在此之前,别的数学家曾试图做此工作但遇到了严重的困难.直到最近,他的思想终于在勒雷的工作中结了硕果,而勒雷再次利用了代数拓扑的工具.

勒雷(Leray, Jean, 1906—),法国数学家.

2 Abelian 函数和代数几何(学)

当庞加莱一接触到黎曼和魏尔斯特拉斯在 Abelian 函数和代数几何(学)方面的工作后,立刻被这领域所吸引.他在这方面的论文在他的全部工作中所占的篇幅同自守函数的工作一样多,发表的日期起自1881年直到1911年.这些论文中的一个主要的思想就是 Abelia 函数的"约化".将雅可比、魏尔斯特拉斯和皮卡所研究的特殊情况加以推广,庞加莱证明了一般的"完全可约性"定理,此定理现在表达如下:若 A 为一 Abelian 族,而 B 为 A 的一个 Abelian 子族,则存在一 A 的 Abelian 子族 C 使得 $A = B + C$,且 $B \cap C$ 为一有限群.这样,Abelian 族可以分解成一些有有限交的"简单"的 Abelian 族的和.庞加莱进一步注意到相应于可约族(甚至椭圆曲线之积,即一维的 Abelian 族)的 Abelian 函数在所有 Abelian 函数中是"稠密"的——这一结果使他能把黎曼的 θ 函数的许多结果加以推广和扩充,并使他能研究相应于代数曲线的雅可比 an 族的 θ 函数的特殊性质.

皮卡(Picard, Charles ÉEmile, 1856—1941),法国数学家.





庞加莱对代数几何最突出贡献是他在1910年到1911年期间的论文,这些论文研究了在代数曲面 $F(x,y,z)=0$ 上的代数曲线.按照皮卡的一般方法,庞加莱考虑了平面 $y=$ 常数和曲面的截线 C_y;除 y 的孤立值外,这样的曲线 C_y 的亏格是常值.这就可能在 C_y 上定义 P 个第一类 Abelian 积分 $u_1,u_2,\cdots,u_p$,而这 P 个积分在这曲面上(更确切地说,在它的万有覆盖上)是解析函数.现在,对这曲面上的每一条和生成曲线 C_y 有 m 个交点的代数曲线 Γ,庞加莱定义了 P 个 y 的函数 $V_1,V_2,\cdots,V_p,V_j(y)$ 是积分 u 在 C_y 同 Γ 的 m 个交点上的值之和;而且,他能利用曲线 Γ 不再出现的那些地方的性质来刻画这些"正规函数",这样,他得到一种代数曲线的解析替代物.这一卓越的方法使他能对皮卡和塞韦里的深奥结果做出简单的证明,同时给出卡斯泰尔诺沃、恩里克斯和塞韦里所叙述的一个著名定理第一个正确的证明.这定理指出了曲面的非正则性 $q=P_g-P_a$(P_g 和 P_a 分别为几何和算术亏格)正好等于在此曲面上的曲线的"连续非线性系统"的最大维数.后来发现,意大利几何学家提出的证明方法是有缺陷的,而直到1965年,除了庞加莱的证明外还没有得到别的证明.他的方法在别的近代问题中也有重大价值(Igusa, Griffths),好像这方法的功用还远远没有穷竭.

3 数　论

庞加莱是埃尔米特的学生,他早期的一些工作是研究型的算术理论中的埃尔米特"连续约化"方法,特别是研究刚由若尔当证明过的(具有非零判别式的)型的类的有限定理.这时期他的一些文章是埃尔米特和若尔当的结果的补充和精确化,而没有引入任何新的思想.与此相关,他给出了具有整系数的型的亏格的第一个一般定义,这是对高斯和爱森斯坦定义的推广;而在此同时,闵科夫斯基独立地获得了这一定义.

庞加莱在数论方面的最后一篇论文(1901)是最有影响的一篇文章,也就是我们现在称为"有理域(或代数数域)上的代数几何"这一理论的第一篇文章.这篇文章的主题就是在曲线 $f(x,y)=0$ 上寻找有理坐标点的丢番图问题,这里 f 的系数是有理数.庞加莱立即证明这问题在双有理变换下是不变的.因此,他自然地去考虑曲线 $f(x,y)=0$ 的亏格,当时他主要考虑亏格为1的情况;利用曲线的椭圆函数参数表达式(或,如我们所说的曲线的雅可比),他注意到,就雅可比而论,有理点相应于一子群,并且他定义这曲线的"秩"就是此子群的秩.庞加莱猜想这秩总是有限的;而这猜想由莫德尔于1922年加以证实,并由韦伊于1929年推广到任意亏格数的曲线上去.他们的证明都是利用了基于椭圆(或Abelian)函数对分的"无穷下降"方法;庞加莱在他的文章中还作了同椭圆函数的三等分有关的类似的计算,而这些思想可能就是莫德尔的证明的起点.

卡斯泰尔诺沃(Castelnuovo, Guido, 1865—1952),意大利数学家.

恩里克斯(Enriques, Federigo, 1871—1946),意大利数学家、哲学家及数学史与科学史家.

塞韦里(Severi, Erancesco, 1879—1961),意大利数学家.

埃尔米特(Hermite, Charles, 1822—1901),法国数学家.

若尔当(Jordan, Marie Ennemond, Camille, 1838—1921),法国数学家.

高斯(Gauss, Carl Friedrich, 1777—1855),德国数学家、物理学家、天文学家.

爱森斯坦(Eisenstein, Ferdinand Gotthold Max, 1823—1852),德国数学家.

闵科夫斯基(Minkowski, Hermann, 1864—1909),德国数学家.

丢番图(Diophantus of Alexandria, 活动于250—275前后),希腊数学家.

雅可比(Jacobi, Carl Gustav Jacob, 1804—1851),德国数学家.

莫德尔(Mordell, Louis Joel, 1888—1972),英国数学家.

韦伊(Weil, André,1906—),法国数学家、数学史家.

克罗内克(Kronecker, Leopold, 1823—1891),德国数学家.

施图迪(Study, Eduard, 1862—1930),德国数学家.

E·嘉当(Cartan, Elie Joseph, 1869—1951),法国数学家.

弗罗贝尼乌斯(Frobenius, Ferdinand Geory, 1849—1917),德国数学家.

韦德伯恩(Wedderburn, Joseph Henry Maclagen, 1882—1948),苏格兰-美国数学家.

阿廷(Artin, Emil, 1898—1962),奥地利数学家.

莫德尔-韦伊定理现在是丢番图方程理论中的基本定理,但关于庞加莱所引入的秩的概念中的许多问题仍未解决,如果对庞加莱的文章作深入的研究可能会得到新的结果.

4 代数学

庞加莱不一定知道克罗内克曾说过这样的名言,他说代数只是数学的女仆,是无权独立地存在的.但无论如何,庞加莱从没有为研究代数这一学科本身而去研究它,只是当他研究算术或分析问题时需要用到代数结果时才去研究的.例如,他对型的算术理论的探讨引导他去研究次数大于等于3的型,而这就容纳了自同构的连续群的研究.看来,超复数系统(在 R 或 C 上)和由这系统可逆元素的乘法所定义的连续群之间的关系之所以吸引他也是同这一问题有关的;在1884年发表的一篇关于这题目的短文激起后来的施图迪和E·嘉当在超复数系上的研究工作.一个很少为人知道的事实就是庞加莱在1903年一篇关于线性微分方程的代数积分的论文中回过头来研究不可交换的代数.他所用的方法使他能引进方程组的群代数(那时是有限的),并(按照H. Maschke的定理,当时显然他并不知道这一定理,而是参考了弗罗贝尼乌斯定理来证明的)将它分离为 C 上的简单代数(亦就是矩阵代数).从而,他首先引进了代数中左理想数和右理想数的概念,并证明了矩阵代数中任何左理想数是最小的左理想数的直接和(一般认为这一结果是归于韦德伯恩或阿廷的).

庞加莱是他那个时代理解和承认李和他的后继者在"连续群"方面的工作的少数数学家之一,特别是在1900年时就体会到E·嘉当文章的深度和广度的唯一的数学家,1899年庞加莱对用新的方法去证明李的第三基本定理,以及现在称做Campban-Hausdorff公式的问题感兴趣,在其工作中庞加莱第一次大体上定义了我们现在称之为李代数(在复数域上)的"包络代数",并给出了从给定的李代数的基推演出包络代数的"自然基"的描述;这定理(更后由伯克霍夫和维特再次发现,现在称为庞加莱-伯克霍夫-维特定理)成为李代数现代理论中的基本定理.

5 微分方程和天体力学

微分方程理论和它在动力学中的应用显然是庞加莱数学思想的中心;从他的第一篇(1878)到最后一篇(1912)的文章中,他从所有可能的角度以全力研究这一理论,很难得有一年没有发表这方面的文章.我们已看到积分具有代数系数的线性微分方程的思想从一开始就支配着自守函数的整个理论.同时,庞加



莱研究了线性微分方程在“非正则”奇点的邻域内的局部问题，首次指出怎样得到积分的渐近展开式，稍后些(1884)他着手对关于 y 和 y' 为代数的并具有固定奇点的所有一阶微分方程(在复域中)的计算，这问题也是由 L·富克斯开始研究的；他的成果由皮卡推广到二阶方程的情况，后来又诱导出班勒卫和他的学派在20世纪初做出的引人注意的结果.

庞加莱最杰出的成果是微分方程的定性理论，也起始于他的创作的全盛时期(1880～1883年)(使我们想起高斯1797～1801年的日记)，这一理论从无到有萌发出来，几乎在其创立者的手中立即就达到完美的程度，这样的数学理论的例子是屈指可数的.庞加莱在1880年和1886年之间发表了四篇有关这理论的长文章，其中尤以前二篇处处闪烁着新的思想.

(1)问题.直到1880年只限于研究微分方程基本类型的解(用“积分法”可求积的)和局部“存在性定理”，而对一般的全局性定理的研究一直局限于线性方程，并且主要限于复数域内(除斯图姆－刘维尔理论以外).庞加莱从一般的方程 $dx/X = dy/Y$ 开始研究，这里 X, Y 为 x, y 的“任意”实多项式，他毫不踌躇地认为可能解最一般的问题，即定性地描述方程的所有解.为了处理无限多条积分曲线，他给出一个好方法，即把(xy)平面投影到与它坐标原点相切的单位球面上去(这个球的球心不在平面上)，这样第一次在一紧致的流形上来研究积分曲线的向量场.

(2)方法.研究的出发点是考虑满足方程 $X = Y = 0$ 的“临界点”.庞加莱应用(加以修改后以考虑到实坐标的限制)柯西和布里奥－布凯提出的著名的“结点”、“鞍点”、“焦点”和“中心”的分类法，为了研究积分曲线的形状，庞加莱引进了“横截”弧的基本概念，弧上任一点都不同向量场相切.此时，这样的一些函数 $F(x, y)$ 的起着主要的作用，即对某一确定的值 C，$F(x, y) = C$ 是横截的(这函数的引入是后来李雅普诺夫在稳定性问题上应用的方法的先驱).

(3)结果.“古典”微分方程的例子使得人们认为“一般的”积分曲线可以由方程 $\Phi(x, y) = C$ 来给出，这里 Φ 是解析的，C 是任意常数.相反，庞加莱指出，这情况仅是在一“例外”情况中成立，即当这些临界点中没有结点也没有焦点时.一般情况是，没有中心——仅有有限个结点、鞍点或焦点；有有限多个闭的积分曲线，其他积分曲线或者联结两个临界点或者“渐近”地趋于这些闭的曲线.最后他指出如何在一些明显的情况下使用他的方法把球剖分成一些区域，这些区域或是不包含闭的积分曲线，或仅包含一条闭的积分曲线.

在这组文章的第三篇中，他着手研究更一般的一阶方程 $F(x, y, y') = 0$，这里 F 是多项式.通过考虑曲面 $F(x, y, z) = 0$，他指出了这问题是确定在一紧

李(Lie, Marius Sophus, 1842—1899)，挪威数学家.

伯克霍夫(Birkhoff, Garrett, 1911—)，美国数学家、逻辑学家.

维特(Witt, Ernst, 1911—)，德国数学家.

L·富克斯(Fuchs, Laszlo, 1924—)，匈牙利数学家.

班勒卫(Painlevé, Paul, 1863—1933)，法国数学家.

斯图姆(Sturm, Charles-Frcm cois, 1803—1855)，瑞士数学家、物理学家、理学家.

刘维尔(Liouville, Joseph, 1809—1882)，法国数学家.

布里奥(Briot, Charles Auguste Albert, 1817—1882)，法国数学家、物理学家.

布凯(Bouquet, Jean-Claude, 1819—1885)，法国数学家.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

致的代数曲面 S 上的向量场的一个特殊情况. 这使他立刻引入 S 的亏格 p 作为这问题的基本不变量,并发现以下关系式

$$N + F - C = 2 - 2p \quad ③$$

这里 N, F, C 相应为结点、焦点和鞍点数. 然后他开始证明他先前关于球的结论可部分地推广到一般情况,而后,对 S 是一环面($p = 1$)的情况作了详细而漂亮的研究,这时可能没有临界点;他遇到了新的情况,要用到关于积分曲线的"各态历经假设". 他没能证明这假设对于一般情况是成立的(在向量场光滑的条件下),但后来由当儒瓦证明了这假设对于一般情况是成立的.

李雅普诺夫(Liapunov, 1857—1918),俄国数学家、力学家.

庞加莱的第四篇论文开创了高阶方程的定性理论,也就是相当于在维数大于等于 3 的流形上的积分曲线的研究. 随着维数的增加临界点的类型也增多,但庞加莱引入了临界点的"克罗内克指数",并证明了包含在以横截超曲面 Σ 为界的有界区域中的临界点的指数和只依赖于 Σ 的贝蒂数后,他发现如何能把在二维情况的关系式③推广到维数大于等于 3 的情况. 而要对所有积分曲线能得到如同在二维的情形下那样精确的描述好像是没有希望的. 可能受他的立体问题的第一个结果(开始于 1883 年)的影响,他仅限于研究一闭积分曲线 C_0 "附近"的积分曲线. 他考虑 C_0 上的一点 M 及在点 M 处和 C_0 正交的超曲面的一小部分 Σ. 如果 Σ 上的一点 P 充分接近 M,那么通过 P 的积分曲线首先在点 $T(P)$ 和 Σ 再次相交,因此可以定义一个 Σ 到自身的变换 T,设 M 保持不变,可以证明这变换是连续可微的(并且如从解析数据出发,则还是解析的). 庞加莱证明了"C_0 附近"积分曲线的特征怎样依赖于在点 M 正切于 T 的线性变换的特征值,因此各种类型的分类非常类似于临界点的分类.

贝蒂(Betti, Enrico, 1823—1892),意大利数学家.

1885 年以后,庞加莱关于微分方程方面的大部分文章是和天体力学,特别是和三体问题有关. 好像巴黎大学的教学使他开始对这方面发生兴趣;在 1885 年时,瑞典国王 Oscar 二世提出在所有国家的数学家中进行一次有关 n 体问题的竞赛. 庞加莱投了一篇长稿,获得了一等奖,这篇文章是他的许多杰作之一,同他的微分方程定性理论的论文有同样的地位. 这篇文章的中心论题是研究三体问题的周期解,这三体问题中的两个的质量相对于第三体是很小的(这情况在太阳系中出现). 在 1878 年 G·W·希尔曾给出这样解的一个例子;1883 年庞加莱巧妙地应用克罗内克指数证明了这样解的整个连续统的存在性. 然后,在他得奖的论文中,他研究了当两个小质量中之一被忽略,在 Hamiltonian 系统中引入参数 μ 的情况,给出了"限制性"三体问题的另一个证明. 由 $\mu = 0$ 时周期解的存在性出发,庞加莱应用柯西的强函数方法证明了对于充分小的 μ 存在周期解的"邻域",他又证明了当 t 趋于 $+\infty$ 或趋于 $-\infty$ 时存在趋近周期解的解,或者甚至当 t 趋于 $+\infty$ 和趋于 $-\infty$ 时都存在渐近于周期解的解("双渐近

G·W·希尔(Hill, George William, 1838—1914),美国数学家、天文学家.





解").我们应强调指出,庞加莱为了得到这样的结果,他首先必须发明一些必要的通用的工具:首先是"变分方程",它给出了微分方程组向量解 f 相对于一参数的导数来作为一线性微分方程的解;其次是相应于 f 是周期的情况下的"特征指数";以及一个向量场的"积分不变量"概念,这推广了刘维尔和玻尔兹曼应用的不变体积的特殊情况.

玻尔兹曼(Boltzmamn, Ludwig Eduard, 1844—1906),奥地利物理学家、数学家.

拉格朗日(Lagrange, Joseph Louis, 1736—1813),法国数学家、力学家、天文学家.

6 天体力学

庞加莱在天体力学方面的工作和他的前辈们的工作形成鲜明的对照.自拉格朗日以来,对太阳系进行数学和数值研究时,是把行星坐标展成按行星或卫星质量(和太阳质量相比是很小的)的幂级数;这级数的系数作为时间 t 的函数,是从由比较运动方程两边质量的同次幂的系数而得到的方程中,用各种近似方法来计算的.开始时,这样定义的 t 的函数不仅包含形如 $\sin(at+b)$(a, b 为常数)的三角函数,还包含了形如 $t\cdot\cos(at+b)$的项等等,当 t 很大时,这些项可能和观察到的运动相矛盾,因此这样的近似方法是不能使人满意的.以后,到了 19 世纪,一些更老练的方法取代了早期的近似方法,这新的级数仅含以 a_nt+b_n 为变量的三角函数;虽然大多数天文学家认为它们是收敛的,但没有人能证明这一点.庞加莱的结果之一就是证明这级数不是一致收敛的,但还是可以用来提供坐标的渐近展开式.

马克劳林(Maclaurin Colin, 1698—1746),英国数学家.

泰特(Tait, Peter Guthrie, 1831—1901),英国物理学家、数学家.

汤姆森(Thomson, William, 1824—1907),英国物理学家、数学家.

因此,庞加莱开创了对天体力学的严格研究,而在他以前的工作都是半经验计算的方法.但他还是热衷于"古典"的计算方法,他发表了近百篇有关太阳系理论方面的文章,在这些文章中他提出了无数的改进和新的技巧.这方面的部分结果收在他的有名的三卷"天体力学方法"及"天体力学讲义"中.从理论观点来看,应提到他的"限制性"三体问题的证明,这里 Hamiltonian 函数依赖于 4 个变量(x_1, x_2, y_1, y_2)及参数 μ,并对这 5 个变量是解析的,对 y_1 和 y_2 是以 2π 为周期的,这时除了 Hamiltonian 函数之外,这运动方程没有具有类似性质的"首次积分".庞加莱还开始了对动力系统"稳定性"的研究,虽然他没有像后来的作者(从李雅普诺夫开始)那么对此概念给出较精确的描述.他证明的最著名的一个结果现在称为"庞加莱递推定理":对"几乎所有的"轨道(对一容许有"正"的积分不变量的动力系统),对一串趋于 $+\infty$ 的时间值,轨道与一任意的非空开集相交.在这定理中特别使人感兴趣的可能是第一次在分析问题中引入了零测集(当然庞加莱没有提到测度,而是说"概率").

庞加莱另一篇著名的天体力学论文发表在 1885 年,是关于仅受重力作用的旋转流体团的形状问题.在这以前,马克劳林发现某些旋转椭球体是这些流体团的可能形状,雅可比指出另一些可能形状是有不等半轴的椭球,泰特和汤

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

贝克勒尔(Antoine Henri, 1852—1908),法国物理学家.
洛伦兹(Lorentz, George, 1910—),荷兰数学家.
爱因斯坦(Einstein, Albert, 1879—1955),德国-美国科学家.
狄利克雷(Dirichlet, Peter Gustav Lejeune, 1805—1859),德国数学家.
施瓦茨(Schwartz, Laurent, 1915—),法国数学家.
C·G·诺伊曼(Neumann, Carl Gottfried, 1832—1925),德国数学家、理论物理学家.
勒贝格(Lebesgue, Henri Léon, 1875—1941),法国数学家.
希尔伯特(Hilbert, David, 1862—1943),德国数学家.
拉普拉斯(Laplace, Pierre-Si-mon, 1749—1827),法国数学家、天文学家.

姆森认为还可能取环状.庞加莱对此问题作了深入的研究,认为还有别的"Pyriform"形状存在.他的论证的一个引人入胜之处是他(显然是第一次)碰到了求"无限"多变量的二次型的极小值问题.

最后,在他较后的(1905)的一篇论文中,第一次研究在一凸的光滑曲面上(他假定是解析的)的闭测地线的存在性这一困难问题.他试图用来证明这种测地线存在性的方法来自于三体问题周期轨道的思想.以后的工作表明,这一方法不够完善,但它启示了许多人,使他们最终成功地得到该定理的完善的证明,并做出了广泛的推广.

7 偏微分方程和数学物理

庞加莱在巴黎大学主讲数学物理有过二十多年;他以他特有的一丝不苟的精神和充沛的精力投入这一工作.结果,实际上他成为理论物理各方面的专家,发表了七十多篇论文和许多不同专题的书,其中他最感兴趣的是光和电磁波的理论.19世纪末,一些新的发现和新的思想革新了物理界,庞加莱有两次在发展这些新思想和新发现中起了重要作用.他对X射线和磷光现象之间可能具有的关系的讨论,正是贝克勒尔实验的出发点,这些实验导致贝克勒尔发现放射性.另一方面,从1899年开始,庞加莱热心于洛伦兹电子理论的讨论;他第一个发现洛伦兹变换构成一个群,这群同构于一使二次型 $x^2+y^2+z^2-t^2$ 不变的群;许多物理学家认为,庞加莱应和洛伦兹和爱因斯坦一起分享发现相对论的功绩.

庞加莱对物理问题的长久兴趣必然使他去研究由数学物理偏微分方程提出的数学问题,而在1880年左右,它们大部分还处于很不成熟的阶段.在他关于这方面的所有论文中,有一典型的特点,就是他对于所用的方法和得到的结果决不忽视其可能有的物理意义(常常反映在物理理论中).这一点在他的最长的一篇关于狄利克雷问题的论文中特别明显.在当时,用C·G·诺伊曼和施瓦茨的两个方法,可得到以曲面 S 作为边界的有界区域 D 内部的解的存在性(对 S 上任意给定的连续函数),但对 S 要加上相当严格的条件.庞加莱给出了第三个方法,即"括去法":这问题等价于 S 上的正质量的存在性,在 D 内这质量的势 V 等于1,且在整个空间是连续的.庞加莱从一包含 D 的且内部势为1的大球 Σ 上的质量着手,然后他注意到由经典的庞加莱公式可以把一个球 C 表面上的质量取代球 C 内的质量,而使 C 外的势能不变,C 内的势减少.用一球序列(C_n)来覆盖 D 的外部,对每一 C_n(以适当的次序),重复应用上面的方法,他指出,当最初在 Σ 上的质量最终被"括到"S 上后,此时得到的势的极限就是问题的解 V.当然,他得证明 V 在 S 的点上有连续性,为证明这一点,他所作的唯





一的假定是,在这些点的每一点,有一以此点为顶点的半锥面(张角$2\alpha>0$),并使得此半锥面同顶点的一邻域之交不在 D 内(稍后,勒贝格的例子说明这样一个限制是不能去掉的).后来,在出现现代的希尔伯特空间方法以前,最初的这方法对势理论的更新起着重要的作用.

在 1890 年同一篇论文中,他开始致力于我们现在称之为拉普拉斯算子的特征值问题的研究,他对此作了长期的努力,但只获得部分的成功.在一些物理问题中(薄膜的振动,固体的冷却,潮汐理论等),遇到如下的问题,即找出函数 u,使它在有界域 D 内满足以下方程

$$\Delta u+\lambda u=0 \quad ④$$

在 D 的边界 S 上满足条件

$$u+k(\mathrm{d}u/\mathrm{d}n)=0 \quad ⑤$$

此处 $\mathrm{d}u/\mathrm{d}n$ 是法向导数,λ 和k 为常数.由富有启发性的变分论证法(狄利克雷原理的黎曼方法的推广)以及模仿斯图姆 - 刘维尔问题(相当于单变量函数的有关问题)可得到如下的猜想:对给定的数,存在一递增的实数序列("特征值")

$$\lambda_1<\lambda_2<\cdots<\lambda_n<\cdots$$

使得问题仅当λ 等于λ_n 之一时可解,且仅有一个解 u_n,使$\int_D u_n^2\mathrm{d}\tau=1$,特征函数 u_n 构成一正交系,对于薄膜振动的例子,u_n 相应于用实验可测出的"谐振动".但在庞加莱之前,λ_n 和u_n 的存在性的严格证明一直没有做出来;对 $k=0$ 的情形,施瓦茨用下面的方法证明过 λ_1 的存在性:模仿斯图姆 - 刘维尔问题可得到,对任意光滑的函数 f,方程

$$\Delta u+\lambda u+f=0 \quad ⑥$$

当 λ 不等于λ_n 时,有唯一的解 $u(\lambda,x)$ 满足 ⑤,此解为 λ 的亚纯函数,以 λ_n 为一单极点.施瓦茨证明了,作为 λ 的函数,解 $u(\lambda,x)$ 等于一有有限收敛半径的幂级数.皮卡还曾证明了 λ_2 的存在性.在 1894 年,庞加莱(还是 $k=0$ 的情形)证明了 $u(\lambda,x)$ 的上述性质,他是巧妙地修改了施瓦茨方法,另外应用了下面类型的不等式

$$\iiint_D V^2\mathrm{d}x\mathrm{d}y\mathrm{d}z\leqslant C\iiint_D\left[\left(\frac{\partial V}{\partial x}\right)^2+\left(\frac{\partial V}{\partial y}\right)^2+\left(\frac{\partial V}{\partial z}\right)^2\right]\mathrm{d}x\mathrm{d}y\mathrm{d}z \quad ⑦$$

(C 为只依赖于D 的常数).这不等式对所有满足 $\iiint_D V^2\mathrm{d}x\mathrm{d}y\mathrm{d}z=0$的光滑函数成立(这不等式是在偏微分方程现代理论中起到基础作用的许多类似不等式的先导).但他没能把他的方法推广到 $k\neq 0$ 的情况,因为找在 S 上有法向导数的 ⑥ 的解是很困难的(他仅能得到我们现在称为"弱" 导数或分布理论意义下的导数).

两年后,当他试图推广诺伊曼的方法求狄利克雷问题的解(仅对凸区域成

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

立)时,遇到了同样的困难.经过对这方法(基于所谓"双层"势)的深入研究,庞加莱把它同上面所说的施瓦茨方法联系起来,这就引出一包含一个参数 λ 的新的"边值问题":找出一个由 S 上的质量所定义的"单层"势 φ,满足 $(\mathrm{d}\varphi/\mathrm{d}n)_i = -\lambda(\mathrm{d}\varphi/\mathrm{d}n)_C$,这里下标 i 和 C 表示法向导数的方向朝 S 内还是朝 S 外.在这里,富有启发的变分论证法使庞加莱再次确信,对此问题有一"特征值"序列和相应的"特征函数",但由于同一原因,他没有能证明它们的存在性.几年之后,他用弗雷德霍姆的积分方程理论解决了所有这些问题;庞加莱的论文对弗雷德霍姆方法的发展似乎起了决定性的影响,特别是在积分方程中引入一个复变参数的思想.还应当说明的是,弗雷德霍姆的行列式是直接受到科克的"无限行列式"理论的启发,而科克的理论本身又是庞加莱关于线性微分方程更早期的结果的一个发展.

弗雷德霍姆(Fredholm, Erik Ivar, 1866—1927),瑞典数学家.

科克(Koch, Helge von, 1870—1924),瑞典数学家.

莱夫谢茨(Lefschetz, Solomon, 1884—1972),美国数学家.

德拉姆(De Rham, Georqes-William, 1903—1969),比利时数学家.

霍奇(Hodge, William Vallance Douglas, 1903—1975),英国数学家.

8 代数拓扑

贯穿庞加莱数学研究工作的一条主线显然是"连续性"观念:我们可看到,无论何时,当他研究分析中的某一问题时,他总要研究,当问题的条件连续变化时,会发生什么结果.因此他每次都必然遇到我们现在称之为拓扑学的问题.他本人在 1991 年说过:"我所致力研究的每一问题都把我引导到位置分析."在微分方程和多重积分的周期的研究中,尤其是这样.从 1894 年开始算起的十年时间内,他写了六篇著名的论文,开创了代数拓扑的近代方法.在此以前,唯一有意义的一步是推广了曲面的"联络的阶"的概念,这一概念是由黎曼和贝蒂独立定义的,庞加莱称之为"贝蒂数"(它们是 $1+h_j$,h_j 是今天的贝蒂数);而实际上除此定义以外没有什么进展.我们现在所称的单纯形下同调的方法完全是庞加莱创立的:包括流形的三角剖分,单纯复形,重心重分,对偶复形,一个复形的关联系数矩阵等概念以及由这矩阵来计算贝蒂数,借助于这些工具,庞加莱把关于多面体的欧拉定理推广了(现在称为欧拉-庞加莱公式),并发现了著名的与流形上下同调有关的对偶定理;稍后,他引入挠率的概念.此外,在他的第一篇论文中定义了流形的基本群(或第一个同伦群),并指出它同最初的贝蒂数的关系.在后面几篇论文中,他给出两个流形有同一下同调,但有不同的基本群的例子.在第一篇论文中他还把贝蒂数同微分形式的积分周期联系起来(他是在研究多重积分和不变积分中熟悉后者的),并且叙述了德拉姆在 1931 年首次证明的定理.人们一直认为直到 1933 年提出更高的同伦群之前,代数拓扑的发展完全是以庞加莱的思想和方法为基础的.

庞加莱为了解决这些问题发明了新工具,另外,他还说明了如何运用这些新的工具到其中的一些问题上去.在一系列关于拓扑学的文章中,有两篇是确



定一代数(复)曲面的贝蒂数,以及确定由方程 $z^2 = F(x,y)$(F 为多项式)所定义的曲面的基本群,这为后来莱夫谢茨和霍奇的推广铺平了道路.在微分方程方面的最后一篇论文中(1912)庞加莱把限制性三体问题(但对参数 μ 不加限制)周期解的存在性问题简化为一个满足某些条件的平面的连续变换的不动点存在性定理,这可能是在代数拓扑学的基础上用分析方法证明存在性的第一个例子.但他没完成这不动点定理的证明,而是在他死后不几个月内,由伯克霍夫获得了证明.

9 数学基础

随着他国际声望的增长,庞加莱越来越多地被要求对广大的读者作数学和科学的各种题目的讲演和书面介绍,对这些琐碎的事他似乎从没有表示不耐烦(1910 年他甚至被要求谈一下彗星对气候的影响问题).他有生动的风格和清晰的思维,是当时一个向外行讲授数学的最好的行家,因而,他的声誉更高了.尽管庞加莱的想象力是变幻不定的,但他对数学发现过程的有名论述至今还是卓绝无比,并且已为许多数学家所证实;他对几何学公理以及它们同实验科学的关系所作的研究是经典的工作,一直有人称他为"哲学家",他是否堪称为一个"哲学家",这一问题最好是让职业的哲学家去决定,而我们只限于看看他的作品对数学基础问题的影响.

庞加莱在物理学上被人们责备为太保守了,但他对待新的数学思想却十分开明.从他在论文中所引用的东西可以发现他读得很广,即使不很系统,但实际上他了解每一个数学分支的所有最新发展.他可能是第一个把康托集合论定理用到分析中的数学家;19 世纪 80 年代初,他在研究自守函数或微分方程时碰到了像完全非稠密集这样一类概念.在某种程度上,他也赞同数学上的公理化倾向,在 19 世纪末,这一倾向还在发展,他对希尔伯特的"几何基础"也表示赞扬.但是,由于庞加莱在 20 世纪早期关于集合论的"悖论"和数学基础争论中的立场,使他成为直觉主义学派的先驱.而这些问题他从没有明白地说过自己的想法,只限于对罗素、皮亚诺和希尔伯特学派的批评.尽管他是接受了数学的"算术化"思想,但他不同意把算术归约为集合论或是归约为皮亚诺的自然数的公理化定义.对庞加莱说来(像以后的布劳威尔)自然数构成最基本的直觉概念,理所当然的是用不着多加分析的;他数次明确地拒绝无限集合的概念而赞成"潜无限",但是他却没有系统地发展这一思想.显然,他对数学的形式化的指责有盲目之处,他常常取笑皮亚诺和罗素学派在这方向上的努力,但他对希尔伯特早期努力的批评意见也许正是以后元数学的发展中某些最丰硕的成果的出发点.庞加莱强调指出,希尔伯特用一个公理系统来表示研究对象的观点,只

康托(Cantor, Georg Ferdinand Ludwig Philipp, 1845—1918),德国数学家,集合论的创始人.

罗素(Russell, Bertrand Arthur william, 1872—1970),英国数学逻辑学家.

皮亚诺(Peano, Giuseppe, 1858—1932),意大利数学家、逻辑学家.

哥德尔(Gödel, Kurt, 1906—1978),奥地利-美国数学家、逻辑学家.

布劳威尔(Brouwer, Luitzen Egbertus Jan, 1881—1966),荷兰数学家.

有当事先证明这一系统不隐含矛盾时才是可取的，现在知道，希尔伯特在 1920 年后建立的理论，即是以无矛盾性的证明为主要目的．庞加莱似乎一直认为这样的努力是徒劳的，而哥德尔的定理证明他是对的；庞加莱未能认识到：元数学领域内的所有这些工作将大大地改进我们对数学推理本质的理解．





庞加莱与米塔-列夫勒[①②]

第四章

瑞典数学家米塔-列夫勒与法国数学家庞加莱在1881年到1911年间保持了一种有规律的通信联系.在瑞典的米塔-列夫勒研究所保留了其中259封信件(米塔-列夫勒所写的119封信件及草稿和140封庞加莱的信件).[③]按照米塔-列夫勒的标准,这个通信量并不算特别大.确实,他几乎像埃尔米特、魏尔斯特拉斯、阿佩尔或班勒卫一样的多产.但是对于庞加莱而言,他与米塔-列夫勒的通信显然是他最重要的通信.还没有搞清楚为什么庞加莱会与他维持这样一种紧密的关系,因为他与米塔-列夫勒并没有数学上的合作.在后文中我将试着解释这一点.

我们可以把他们的通信分为四个时期.第一个时期(从1881年4月到8月的10封信)是这两位数学家开始相互认识的时期.第二个时期的主题是创建《数学学报》(*Acta Mathematica*)和庞加莱关于富克斯函数的文章(从1881年8月到1887年3月间的大约50封信).第三个时期则专注于庞加莱参加为庆祝奥斯卡二世皇帝的六十寿辰举办的竞赛(1887年3月到1890年7月间的大约50封信).在最后一个时期中,我们可以看到两位卓越的数学家在处理大学和科学院里的任命和提职以及诺贝尔奖的提名工作(1891年5月到1911年9月间的大约150封信).

① 原题:The Poincaré-Mittag-Leffler Relationship 译自:The Math. Intelligencar, Vol. 21. No. 2, 1999: 58-63. 许秀兰,胥鸣伟,译. 袁向东,校.

② 本文作者 Philippe Nabonnand 在法国 Nancy 大学教授数学、逻辑和数学史,也在 Elie Cartan 研究所(研究数学课题)和庞加莱档案馆(研究历史课题)工作,是庞加莱通信集的编委.

③ 庞加莱的全部通信将由庞加莱档案馆编撰.第一卷是加注的庞加莱与米塔-列夫勒通信,并有评论.

1 接　触

1881年米塔－列夫勒就“缺项函数”问题写信给庞加莱；所谓缺项函数是指不能拓展到整个复平面的复变函数.之前，埃尔米特已向他谈到过他对庞加莱极高的评价，而米塔－列夫勒要了解庞加莱论文中有关结果的更多的信息以及他在微分方程方面的工作：

> 你想在何时发表你在微分方程方面的研究结果？我等得不耐烦了.从埃尔·米特的报道中我看不出你的结果是否与富克斯刚刚发表的一样，也看不出你的研究更为一般.（米塔－列夫勒致庞加莱，1881年4月11日——IML）

那时庞加莱还没有出名：他只发表过一篇关于微分方程定性理论的通讯和两篇关于富克斯函数的通讯.另一方面，米塔－列夫勒已经是知名的数学家了.他是魏尔斯特拉斯的最有才华学生中的一个，他还对魏尔斯特拉斯的结果——具事先给定的点为零点的全纯函数的存在性定理（魏尔斯特拉斯，1876）作了精彩的推广，他的定理是关于具有给定了极点和零点的半纯函数的存在性的（米塔－列夫勒，1879）.因此埃尔米特对米塔－列夫勒以庞加莱的保护人自居的态度并不感到奇怪①.另外，在他的文章的第一次发表时（庞加莱，1881a），庞加莱没有引述魏尔斯特拉斯关于解析函数的结果，特别是对那些有“缺项空间”的函数；为此，米塔－列夫勒“坦率而真诚”地告诉庞加莱，他必须加以引述.（米塔－列夫勒致庞加莱，1881年5月22日——IML）

米塔－列夫勒一直在坚定地维护着魏尔斯特拉斯的全部工作.随岁月流逝他越加担心魏尔斯特拉斯研究成果的优先权问题，或更一般地，担心他的工作的流失，因为没有发表他全部的成果，其中一大部分仅仅在他的讲课中出现.因此，许多东西只有他的学生才知道而法国的数学家们并没有注意和意识到这点.米塔－列夫勒在许多场合都写信给庞加莱谈到魏尔斯特拉斯的优先权问题.

米塔－列夫勒是严谨的柏林学派的一名高手，他认同魏尔斯特拉斯的观点，因而常常对庞加莱的数学写作提出批评：

① 我想，庞加莱还是个年轻人.（米塔－列夫勒给埃尔米特的信，1881年4月6日——AS）



你是怎样看待庞加莱的第二篇论文《关于富克斯函数》(*Sur Les fonctions fuchsiennes*)的？实在遗憾，他不是德国大学的毕业生．正如他的文章充满了新思想一样，对我而言，他似乎也留给我对它们的严格的表达形式以太多的期待．(米塔－列夫勒致魏尔斯特拉斯，1883年5月11日——IML)

米塔－列夫勒对于庞加莱的工作总持有一种模糊和矛盾的态度．一方面他赞叹庞加莱的“天赋”，而另一方面他也批评他缺少严谨性．

但是他(指庞加莱)……写得太不用心了，他的论文充满了不准确性．这些是只能在你我之间谈论的事情！人们必须让这些天才们按他自己的道路去发展并以感激的心情去接受他们赋予我们的东西，尽管人们可能会希望以更容易消化的形式得到它．(米塔－列夫勒致埃尔米特，1887年10月27日——IML)

2 创建数学学报

虽然如此，在1881年由于对关于富克斯函数文章的深刻印象和埃尔米特的力辩，米塔－列夫勒明白了庞加莱是一个真正的数学天才：

我衷心祝贺你在你的研究工作中取得的巨大成功，我觉得我们亲爱的大师埃尔米特在写给我的信中说你是一个名副其实的数学天才，这是完全正确的．我只希望你把你所有的在富克斯函数方面的研究者汇集起来出版一个大集子．(米塔－列夫勒致庞加莱，1881年6月22日——IML)

出版庞加莱重要文章的机会帮助了米塔－列夫勒做出决定来着手实现创建一份新的数学杂志的计划．在19世纪70年代，法国数学杂志，像 *Journal de mathématiques pureset appliquées* 和 *Journal de l' Ecole Polytechnique* 的水平已经下滑(Gispert，1996)．而在同一时期，德国的数学杂志，像 *Journal für die reine und angewandte Mathematik*(*Crelle* 杂志)或新创建的 *Mathematische Annalen*，反映出德国数学界的巨大活力和创造性．但这些杂志在使用德语的人的自然疆界外并没有得到广泛阅读，特别是在法国．米塔－列夫勒考虑到了这点，想要创立一份新杂志以在德国和法国的数学家之间架起一座桥梁．加之由于他与埃尔米特和魏尔斯特拉斯的友谊，他期待着这两个数学群体之间的合作．

但光有有利的环境还是不够的．米塔－列夫勒记得，*Crelle* 杂志的成功是建

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

立在阿贝尔有关椭圆函数的重要贡献这个坚实的基础上的.现在他争辩说,富克斯函数新理论的传播需要一份新杂志:

> 我向你保证,你的这些新发现足以与阿贝尔的那些相媲美,而你的函数是自椭圆函数发现以来最出色的(……).现在我对你有个建议.我们,斯堪的那维亚的数学家们,有一个计划,要按照 *Crelle* 杂志的模式出一份新的数学杂志,(……)现在我们,Gyldén 先生和我认为你,一位法国人,必将慷慨地确保我们杂志的成功.您同意把你的文稿"关于富克斯群"(Sur les groupes fuchsiens)交由我们杂志的第一期发表吗?(……)我请求你对我们这一计划保守秘密,因为此计划能否实现依赖于你.如果你拒绝,我的看法是,我们不得不再等待两三年.唯有发表你的发现这个有利条件才能说服我立即行动.(米塔-列夫勒致庞加莱,1882 年 3 月 29 日——IML)

庞加莱同意在这份新杂志 *Acta Mathematica*(《数学学报》)上发表他的关于富克斯函数的五篇大作.庞加莱答应帮助米塔-列夫勒的那封信已经丢失,我们只能猜度为什么庞加莱决定在 *Acta Mathematica* 上发表他的文章.首先,我们能够推想庞加莱对于米塔-列夫勒的提议觉得很满足.另外,庞加莱确信他必须迅速行动来确立自己的优先权并提高他的声望,必须使他的理论在德国这个在数学上卓越的国家里为人知晓.还有,我们不要忘记在克莱因与庞加莱之间发生的争吵,它涉及对富克斯函数的优先权和术语的选取.庞加莱很注意使他的结果为某些杰出数学家知晓:

> 一般来说,作者们急不可耐地想得到他们的抽印本而不急于把它们分发给朋友们,只是把这些抽印本以尽快的速度寄给十来个大名鼎鼎的人物,以期能使他们知道这些工作.(庞加莱致恩内斯特勒姆,1884 年 6 月 3 日——CHS)
>
> 因此,庞加莱需要在德国传播他的工作.他已在 *Mathematishe Annalen* 这份克莱因的刊物上发表了一篇关于富克斯函数的综述文章,但是由于 1870 年战争之后在法国的仇德情绪的缘故,我们猜想他不可能在一份德国刊物上发表他的主要论文,而米塔-列夫勒的提议能让德国人读到他的文章,而又不会冒犯他同胞的民族主义情绪,文章发表在 *Acta Mathematica* 上可同时传到法国和德国的读者手中,对他在这两个国家中的声誉有利.如果说米塔-列夫勒把 *Acta Mathematica* 的成功建立在庞加莱的参与之上,那么庞加莱则把他的国际声誉寄希





望于在那份刊物上发表许多文章.

庞加莱在 *Acta Mathematica* 头 10 期上发表了 10 篇文章(1594 页中有 676 页为法国作者所写).此时正处于庞加莱特别具有创造力和多产的时期,所以这些文章并不全都是关于富克斯函数的.例如,1883 年他证明了两个变量的半纯函数是两个全纯函数的商:

长期以来我一直想弄清楚是否一个半纯函数 $F(x,y)$ 可以写成 $G(x,y)/G_1(x,y)$ 的形式,但没有得到满意的结果.我知道魏尔斯特拉斯先生把这个问题看做是分析中最本质最困难的一个问题.你告诉我你正在研究解决此问题.我希望当你有了一个肯定的结果时就立刻告诉我而不要延误.(米塔-列夫勒致庞加莱,1882 年 12 月 5 日——IML)

另外一些重要文章是"流体的平衡态"(1885a)及"线性方程的不规则积分"(1885b);这后一篇文章是他与 Thomé 发生激烈争辩的起因.这些数学的争端几乎是使庞加莱打破他矜持的唯一原因.我已经提到过他与克莱因的非常激烈的争辩.米塔-列夫勒在若干场合提到过这次争吵,像在 1882 年 7 月 18 日,谈到他拜访施瓦茨的事以及施瓦茨的狂怒:

对施瓦茨先生它可不是同一回事.我发现他对你充满了轻蔑.他认为是他头一个发现了这些不能在椭圆函数论中找到的,被你称为富克斯群的例子.(……)依你的观点,在这场出现于德国数学家中间关于你起的新名字的纷乱争吵中,你必定已找到你的发现的重要性的证明.施瓦茨先生没有隐瞒这样的事实:富克斯函数是使他狂怒的原因.(IML)

在做出回答之前,庞加莱很快地研究了施瓦茨的文章并得出了结论,即他看不出有更改他想法的理由:

我不想抚慰施瓦茨先生.引起他狂怒的理由是什么? 首先,他之所以愤怒是因为虽然他掌握着一个重要结果但却不知道怎样利用它.对此我也无能为力.

其次,他不满意采用富克斯这个名字,他想用施瓦茨冠名.我已说过我做不到这些的理由了.(庞加莱致米塔-列夫勒, 1882 年 7 月 27 日——IML)

庞加莱也以同样的方式回答了魏尔斯特拉斯的关于他的文稿"微分方程的

积分"所提的问题.在这篇文稿中,庞加莱证明总存在一个变量变换使得一个微分方程的解可用一个级数来表示,而此级数在新变量的所有值上均收敛.魏尔斯特拉斯的异议涉及三体问题中的碰撞情形.庞加莱回答说,如果发生碰撞,他的新变量趋向于无穷,因此,在碰撞之后"这些公式没有给出任何东西(……),而那正是它们的最好结果".庞加莱对这位年迈大师的回答外表看来有点漫不经心……

后来,在他取得奥斯卡二世奖的竞赛成功之后,庞加莱在与瑞典天文学家 Gyldén 的辩论中,米塔－列夫勒又帮了他的忙.在他的得奖论文中,庞加莱讨论了这位天文学家运用的新展开级数(没有长期项)的收敛性.

> 当代的几何学家致力于把这些展开换做一些新的展开式,它们只含有三角函数项.最近他们取得了成功,Gyldén 先生的或 Lindstedt 先生的级数仅含的项都具形式 $A\ \sin\ at$ 或 $B\ \cos\ at$.
>
> (……)但是,这还不是最后一步.我们可以问是否这些级数都收敛,而且因为"小因子"的出现会导致一些非常大的项出现,那么收敛性便值得怀疑.这个工作将证明这些怀疑是有根据的:所有这些级数都是发散的.我不得不排除掉 Gyldén 先生在他最近文章中所提出的那些级数.我确实没有任何办法知道它们是收敛还是发散.(庞加莱,1888,6 页)

在没有读过庞加莱的文章情况下,Gyldén 继续认定他自己的展开式的收敛性并宣称他在涉及渐近解的某些结果上比庞加莱领先.奥斯卡皇帝需要从米塔－列夫勒那里得到一个回应,所以米塔－列夫勒要求庞加莱帮助.在他的回答中,庞加莱首先说,阅读 Gyldén 的文章(1887)并要了解他的结果是非常艰难的.他也解释了他自己的阅读方法:

> 我要向你承认,我认为 Gyldén 先生的风格有点乏味而且对我来说读起它来非常困难,对吗?当我阅读一篇论文时首先会很快地浏览一遍以得到一个一般的印象,然后再返回到那些对我来说有些晦涩难懂的地方.我发现自己去做证明比彻底地去检验这位作者的证明要来得更方便.我的证明一般来说相当粗略,但它们对我的好处是它们是我的.(庞加莱致米塔－列夫勒,1889 年 2 月 5 日——IML)

由于 Gyldén 的算法没有准确地定义,要确定它的收敛性是困难的.庞加莱指出 Gyldén 的论证是不严谨的,并不留情面地下了结论:





(Gyldén 的证明)等于接受了下述原理:除非我们有非常严格的理由怀疑,否则每个级数当它的变量小于 1 时都收敛.(庞加莱致米塔 - 列夫勒, 1889 年 3 月 1 日——IML)

庞加莱在同一封信的后面部分承认,Gyldén 的收敛性论证对于天文学中的计算是足够的,但"对几何学家是不够的".

这些数学的争端是少数几个场合能使庞加莱显露出一些豪爽和开朗.在他的大多数通信中他总是非常拘谨的,不表露出任何个人情感.他极少提及政治和社会事件.另外,虽然他对科学院和大学生活感兴趣,但与米塔 - 列夫勒不一样,他从不跟科学界里的闲话和流言沾边.

3 奥斯卡二世奖

由于有了 J. Barraw-Green 的书(1997),奥斯卡二世奖的竞赛情况现已知道得很清楚了.奥斯卡二世是瑞典皇帝也是科学界的朋友.值他六十寿辰之际,米塔 - 列夫勒组织了一场数学竞赛,还成立了一个由埃尔·米特、魏尔斯特拉斯和米塔 - 列夫勒组成的委员会,他们提出的四个问题"从不同的观点同等地吸引那些善于分析的人的注意力,而他们的解答会对科学的进步有极大的好处"(奥斯卡竞赛的通告).第一个问题由魏尔斯特拉斯提出,与 n 体问题有关,特别与轨道的稳定性有关.庞加莱决定参赛:

我没有忘记奥斯卡皇帝的奖,我甚至会说,这一两个月以来我一直想的只是这个奖.(庞加莱致米塔 - 列夫勒,1887 年 7 月 16 日——IML)

庞加莱试图回答第一个问题.他的主要结果是关于限制性 3 体问题的稳定性证明.庞加莱很快就得到了这个结果并希望得到更一般的结果:

不过,我已得到一些有趣的结果并且我要引述其中的一个.它涉及的特殊情形是第一和第二个天体具有有限质量而第三个是零质量.第一个和第二个天体绕着它们共有的重心作圆周运动而第三个则在这些圆的所在平面上运动.在这种特殊情形下我找到稳定性的一个严格证明,并且找到一个方法来确定第三个天体的某些参数的精确的界限.(……)我现在希望就一般情形发起冲击,在 6 月 1 日前如果没有完全解决此问题(我不希望如此)也至少要找到一些结果能足以达到

可送交给竞赛的要求.(同上)

庞加莱围绕这个结果组织了他的论文.在前言中他表达了他的失望:

> 本文着手回答竞赛四问题中的第一个;但我所得到的结果却如此不完全,以至于如果我不知道这个问题的重要性和困难程度会引起人们对与它有关的任何问题的兴趣,而且我们只能期望在长期不断的努力之后才能得到确定的回答的话,那么我就不愿意发表它.(……)我不得不局限于一种特殊情形.我仅处理了在两个自由度下的运动方程.(庞加莱,1888,5~8页)

在1888年夏季期间,米塔-列夫勒和魏尔斯特拉斯研究了庞加莱的论文.米塔-列夫勒非常热情,但是他认为庞加莱的工作非常难,而且某些证明还不完全.庞加莱以100页补充注解回答了他的问题.1889年1月21日奥斯卡皇帝向庞加莱颁了奖.

但是到了七月初,米塔-列夫勒转交给庞加莱一些问题,它们是由*Acta Mathematia*杂志社他的助手Phragmén提出来的.这些问题涉及由庞加莱引进的渐近解展开的收敛性.庞加莱的回答成为最后一篇补充注解"关于渐近解"(Sur les solution aeymptotiques,1888,251~256页).似乎这些问题刺激了庞加莱更精确地去审查渐近解收敛性的所有证明,而且不幸地发现了一处无法修补的错误.在12月1日,庞加莱对米塔-列夫勒绝望地宣布,他文章的大部分都变成是有纰漏的了.

> 我不愿向你隐瞒这次发现带给我的痛苦.首要的是,我不知道是否您仍然认为这个结果依旧(……)应该得到你给予它们的巨大奖赏.(庞加莱致米塔-列夫勒,1889年12月1日——MLJ)

庞加莱的稳定性证明依据的事实是渐近曲面(即由渐近解生成的曲面)是闭的,因而可以排除掉其他的解①.它是由一个更一般的引理推出来的,但这个引理却是错的.

我们可以想象,在一段短时间内庞加莱必定想直接地由渐近解收敛性来证明渐近曲面的闭性,以挽救他的稳定性成果.确实,在12月10日埃尔米特写信给米塔-列夫勒说,这个错误并不像庞加莱所想的那么严重,它不过是个重新

① 如今,渐近曲面被称之为"稳定与非稳定流形".



组织一下他的"令人羡慕的工作"的问题而已.但是庞加莱发现这些展开式只是渐近而不收敛.

> 不幸,这些级数并不收敛(……)但是即使是发散的,难道我们就不能运用它们吗?(……)我们可以说在这命题中所得到的级数在小的 μ 值情形下表示了渐近解,正如同斯特灵公式表示了欧拉函数那样.(庞加莱,1890,384~386 页)

虽然许多结果仍然有效,特别是 Lindstedt 级数的发散性及递推定理,然而它仍是个灾难.米塔-列夫勒保持着冷静,回答说他并不懊悔颁奖给庞加莱.他仍以通常的效率忙碌着.那时,庞加莱的论文已经印好,一部分已递交给 *Acta Mathematica* 的编委们,于是不申明任何理由地收回这些文本便是最紧急的问题.当做完这件事之后,米塔-列夫勒建议庞加莱写出一篇新的论文来发表.

> 那么,现在我来谈建议你做的事.依我看来,它将是对你对我们最为体面的了.你写一篇新的论文,把你原来文章中还留得下来的内容放进去,再把它们展开,然后再加进去任何你认为对的东西.你要为这篇新论文写个前言,说明它是你得奖论文的修改稿,说你的一些新发展仅在这里有所描述,同时还纠正了一个错误(……)我想这件事在你的论文发表前必须不能让别人知道.(米塔-列夫勒致庞加莱,1898年12月5日——IML)

尽管魏尔斯特拉斯抱怨他已在闲言碎语中听到了庞加莱的错误的消息,米塔-列夫勒还是非常成功地保守住了整个故事的秘密.12月20日,庞加莱接受了米塔-列夫勒的交易条件①.之后,米塔-列夫勒认为从公共机构的观点看,危机已经处理妥当了.在万无一失之后,他又怂恿埃尔米特在法国科学院表扬庞加莱的工作.

> 在科学院的正式会议上你必须无条件地谈到他的论文.否则,每个人都会惊讶并开始怀疑你为什么不讲到它.进一步说,你实在是可以赞扬这篇论文而不用担心会夸大其词,因为它仍然是我们这个年代的杰作之一.(米塔-列夫勒致埃尔米特,1889年12月20日——AS)

---

① 特别地,庞加莱同意支付新的印刷费(3 585 瑞典克朗).注意,奖金额为 2 500 克朗而米塔-列夫勒的年薪为 7 000 克朗(Domar,1982).

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

他还说：

> 这个忠告对庞加莱先生将是非常有用的，他或许会在未来的时日放弃他那令人遗憾的习惯：讲述他明知其证明并不完全的结果.(……)当然，我赞同你对庞加莱的看法，承认他的天才，但他给每个人增加的负担太多了；一个人如果不向更加严谨的要求俯就屈从，就不应当保有像他已有的这样一个显要地位.他仍然年轻，他会改进，数学科学将有所收获.(同上)

4 诺贝尔物理奖

庞加莱与米塔-列夫勒关系中最为有趣之处是米塔-列夫勒试图为庞加莱得到诺贝尔物理奖而作的努力.米塔-列夫勒的比较大的目标是要使诺贝尔物理奖能授予从事理论研究的学者.这是他与瑞典大学一群人又一轮的战斗，他们是“Ostwald 和 Arrhenius[①]的科学方法”的支持者.Crawford 在她的书《诺贝尔机构的开端》(*The Beginning of the Nobel Institution*, 1984)中指出，就诺贝尔物理奖来说，在实验工作者和理论工作者的对立中，没有比物理领域的定义更具有利害关系了.

米塔-列夫勒向庞加莱解释了他的战略：

> 现在我正在尽全力为洛伦兹争得诺贝尔奖.把洛伦兹放在庞加莱的前面是必要的.首先，洛伦兹更直截了当地是个物理学家，而且他必须有一个由具有足够资格的权威签署的报告.庞加莱正在起草这份报告.如果我能取得成功，(……)我必定会成功地打开通向理论之门，这必定符合诺贝尔的想法；而后，先是庞加莱，再有你将随之而来.(米塔-列夫勒致 Painlevé, 1902 年 7 月 18 日——IML)

但是，米塔-列夫勒不得不同意妥协，接受洛伦兹与 Zeeman 共享这个奖，后者却是个实验工作者.在 1904 年与 1909 年间由达布发起，庞加莱多次获诺贝尔物理奖提名.在 1910 年，米塔-列夫勒认为当时环境有利.为了协调法国与瑞典两方面所作的努力，他写信给阿佩尔：

① Ostwald 在莱比锡工作，获 1909 年诺贝尔化学奖，Arrhenius 在斯德哥尔摩工作，获 1903 年诺贝尔化学奖.



> 我们可望使庞加莱赢得诺贝尔奖的时机来了.在下封信中我附有由弗雷德霍姆写的一份建议,他信服您的判断;还有一份由达布写的建议.弗雷德霍姆很好地利用了达布今年所提的建议.最重要的事是先确立纯理论在物理中所起的突出作用,然后作为结论可建议将诺贝尔奖给予那些能用十分简洁公式来确定的发现.在一些讨论之后,我们在庞加莱关于数学物理的微分方程方面的发现里找到了这种公式.我想我们按这一程序运作定能够赢得这次竞赛.(米塔-列夫勒致阿佩尔,1909年11月28日——IML)

米塔-列夫勒还说,提名者必须避免"数学"而要提及"纯理论",因为"像那些仅仅是实验物理学家的物理学诺贝尔委员会的成员对数学有一种愚蠢的恐惧感".

阿佩尔,达布和弗雷德霍姆签署了一份报告"关于庞加莱在物理方面的工作(Sur les travaux d'ordre physique de Poincaré)".在前言中,他们声称不能忽视理论的进展在科学前进中所起的非常重要的作用.他们强调说,物理学家必须提出理论并建立数学工具:

> 物理学要达到了解事物隐藏的和内在的机理的水平,假说的作用就要增大;要达到这样的水平,物理学家对更好的分析工具的需求也要增大.(Report on "les travaux d'ordre physique de M. Poincaré"——CHS)

由于他对许多数学物理问题做出了广泛和正确的解答,庞加莱的提名被证明具有充分的理由.而后米塔-列夫勒向整个物理界的学者逐个发出信件要求他们提名庞加莱为诺贝尔奖的候选人.34个著名的物理学家和数学家支持了这份赞成庞加莱的提案①.不幸,实验物理学家和 Arrhenius 圈子里的人没有被说服而诺贝尔委员会把奖授予了一个实验物理学家 Van der Waals.

> 在这次争取诺贝尔奖的活动中,我们又一次被打败了.一大群博物学家都投票反对我们,他们对事物的基本原理一窍不通.他们害怕数学,因为对它的任何东西他们连弄明白一点点的可能性都没有.(米塔-列夫勒致庞加莱,1910年12月6日——IML)

---

① (见 Crawford 1987)但是,某些物理学家像 Rutherford 并不同意庞加莱的提名,因为他们认为庞加莱并不是真正的物理学家.

档案目录

AS——Archives of the Académie des Sciences-Paris(France).

CHS——Center for the History of Sciences—The royal Swedish Academy of Sciences—Stockholm(Sweden).

IML——Mittag-Leffer institute—Djursholm(Sweden).





法国在数学发展中所起的作用①②

第五章

1 优秀的传统

像其他科学一样，数学是国际上的一项公共财产. 它是属于所有国家的共同体，每个国家根据各自的能力对此做出相应的贡献. 一个备受尊敬的数学家，如果对外国昔日伟大的头脑(只提最重要的几位：意大利的伽利略，英国的牛顿，瑞士的欧拉，挪威的阿贝尔，德国的莱布尼兹，高斯和黎曼)不怀有极大的敬意将是难以接受的. 他们在数学的不同领域开辟了新的方向，如果没有他们，数学将不会是今天这个样子. 然而，我想告诉你们法国数学家是对数学的发展做出了最显著贡献的一支力量，而且，当提及伟大数学家的数目时，法国也不会少于任何其他国家. 我很荣幸有这样的机会在如此友善的听众面前、在带给我如此多美好回忆的国度里谈论这个特别的主题.

数学和其他科学一样，存在着两类科学家：第一类通过提出新想法开辟康庄大道，这些想法通常简单但未被他人发现；第二类则在第一类为了到达他们自己的花园而清除过的辽阔土地上劳作，经常可以拣到可口的果实，而且有时还能获得巨大的收获. 当涉及任何科学的发展，后者不仅仅是重要的那么简单，应该是不可或缺的；然而，显然前者更应被人们记住和授予荣誉. 我今天要谈论的就是这些人.

埃利·嘉当(Elie Cartan). 著名法国数学家. 1869年4月9日生于伊泽尔的多洛米约，1951年5月6日在巴黎去世. 埃利·嘉当在李群、微分方程、微分几何等领域做出了开创性的贡献，对20世纪数学的发展有着重大影响，被誉为“现代微分几何之父”. 他的儿子H·嘉当也是著名数学家，曾获得沃尔夫奖.

① 原文是埃利·嘉当于1940年2月27日在南斯拉夫贝尔格莱德的法国研究院作的演讲.
② 原载自：丘成桐，刘克峰，季理真. 数学与生活[M]. 杭州：浙江大学出版社，2007：71-90.

约瑟夫·贝特朗(Bertrand, Toseph Louis Francois, 1822—1900),法国数学家.

约瑟夫·贝特朗告诉我们,国王亨利四世有次在枫丹白露接见荷兰大使,因提起法国人在文学领域中的成就超过他们的外国对手而高兴."那些我自己也崇敬的是数学家."荷兰人——他培养了一个研究几何的数学家——说,"但我注意到,到目前为止,法国还没能产生任何数学家.""你弄错了!(Romanus se trompe!)"亨利四世喊道,同时马上转向他的一个仆人,要求召入 M. de la Bigottere,他是法国第一位伟大的数学家,是现代代数的创始人,真名是弗朗索瓦·韦达.韦达第一个意识到,求解特殊数值方程的过程可以通过对字母作符号运算(其源头可以追溯到古代)得到简化;他同时系统地发展了这个想法,预测它的应用范围可以无限地扩展.16 世纪晚期,当伽利略和高等几何学校使意大利闻名于世之时,正是弗朗索瓦·韦达确保法国在创建现代数学的过程中占据了一个显著的位置.我该告诉你们,韦达和你们国家最早的数学家之一——盖塔尔迪有过相当长的接触.盖塔尔迪出生于杜布罗夫尼克,他于 1600 年在巴黎出版了韦达最后一批著作中的一本.

韦达(Vieta, Francois, 1540—1603),法国数学家.

盖塔尔迪(Ghetaldi, Marino, 1568—1626),南斯拉夫数学家.



韦达



笛卡儿

笛卡儿(Descartes, Renédu perron, 1596—1650),法国哲学家、数学家、物理学家和生理学家.

帕斯卡(Pascal, Blaise, 1623—1662),法国数学家.

费马(Fermat, Pierre de, 1601—1665),法国数学家.

17 世纪的法国尤其辉煌.在数学、力学和物理的历史中,这个时期出现了三位特别著名的人物:笛卡儿、帕斯卡和费马.

哲学家、数学家兼物理学家笛卡儿常被视为人类思想史上一个新纪元的开创者.作为一个物理学家,他在探求真理时出现了一个瑕疵;然而,他关于所有的物理现象都可以表示为空间和运动的想法至今仍充满着魅力,因为广义相对论的创始人自己也相信通过几何概念可以解释物理的可能性(只是数学随后的发展使得爱因斯坦能够比笛卡儿更好地实现他的想法).即使我们不承认他是解析几何的创立者(1637),也不能低估他在数学中所起的作用.我们知道,希腊几何学家已经在他们的思考中自由地使用数字和计算,但数字并没有完全失掉其几何特性,那是他们希腊式的科学中所拥有的:像"square","cube"既指数又

指几何形状,显然,现在日常讲话中仍有这种双关用法.笛卡儿最早系统地使用抽象的数代表几何形状并将几何推理转换成计算.通过这种方式,他创造了一种特别强大的工具.我们应当将这种主要源于解析几何和微分几何的几何学的发展归功于他.他还发现一种寻找代数曲线切线的一般方法,由此丰富了微分几何.由于有了解析几何,数学家不但成功地理解了任意维数的空间,而且知道怎样在这样的空间中进行几何化的思考.可以这么说,正因为解析几何让数学家感到自然,比如三维球面,直到现在物理学家才开始用它来解释物理现象.所有这些,虽然遥远,但毫无疑问是笛卡儿思想和结果的推论.在代数学中,我们将关于符号运算的规则归功于他.在纯几何中,我们要将被欧拉独立发现、现以欧拉的名字命名的一个定理归功于他.这是拓扑学(当时还没有这门科学)中的一个结果,它建立了凸多面体顶点数、边数和面数的关系.最后,力学中关于线动量守恒的笛卡儿原理,直观地解释了(只需适当的精化就能推出)经典力学中的一条基本原理.



帕斯卡



费马

巴雷斯·帕斯卡是个有些怪异但非凡的天才.早在16岁,他就通过写作《圆锥曲线论》(*Traitesur les sections coniques*),显示出不同寻常的几何天赋.《圆锥曲线论》通常研究平面与圆锥面相交所得的曲线,在开普勒的行星运行规律中有着重要的应用.帕斯卡用了他同时代数学家迪沙格的成果.迪沙格是法国最重要的几何学家之一,和帕斯卡一样,是射影几何的先驱.类似于迪沙格的方法,以透视图作为出发点,帕斯卡成功地将圆锥曲线的所有性质化约为一个他叫做"神秘六卦"(L' hexagramme mystique)的性质,即圆锥曲线内接六边形的三对边交点共线.单凭这个结果帕斯卡就展现了他作为著名几何学家的创造力.

迪沙格(Desargues, Girard, 1591—1661),法国数学家.

帕斯卡一确立射影几何先驱的地位,就开始以概率论缔造者的身份登台.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

当他的朋友 Chevalier de Mere 向他询问一些关于赌博的问题时，帕斯卡通过将所有可能出现的结果化约到最基本的情形来作答. 另一方面，皮埃尔·德·费马用完全不同的方法得到同样的结论. 从帕斯卡和费马的信件交往可以看出数学概率论原理是怎么发展起来的. 这个新研究领域没有被帕斯卡遗漏. 他有一个广为人知的叙述："通过结合确定的数学方法和不确定的几率，这个新科学可确切地给予一个令人惊讶的名字——几率的几何."从著名的打赌证明，我们知道他的研究和思维在多大程度上受他对这门新几何兴趣的影响. 我们还知道这种几何是现代科学发展不可或缺的工具，其中整个物理只是数学概率论的章节，而且许多物理定律只是概率的定律.

前面提到过的费马是最伟大的数学天才之一. 他 30 岁时成为参议员且终身保留着这个位置. 虽然他的职业不会带给他数学上的名望，但是他确实为自己所钟爱的事业花费了许多时间. 费马尤其因在算术和数论方面的研究而著名. 他也是《趣味问题集》(*Problemes plaisants et delectables*)的作者. 关于不定方程部分的空沿上，他写下许多重要但未证明的定理，通常认为他知道它们的证明. 其中最著名的要数通常被称为费马大定理的那个，根据这个定理，两个整数的 n 次方之和不会等于第三个整数的 n 次方，其中 n 是任意大于 2 的整数. 这个定理激发了大量的结果，尽管它们的作者手头拥有费马那时所不知道的现代代数结果，但还是不能证明它或找出反例. 长久以来，大家一直认为，即使这个定理总体是错误的，但其实可能只是对某些特殊的 n 是错误的；然而，无论如何都不知道某些导致定理不成立的 n 究竟是有限的还是无限的. 单凭这个定理引发的研究——贯穿在几乎所有由数学发展起来的理论——费马影响了数论的发展. 他的同代人能立即认出他在那个领域的不同寻常的技巧. 在一封信中，帕斯卡写到他在数论中的结果被费马的所超越，他只能对它们充满敬意.

卡瓦利里(Cavalieri, Bonaventura, 1598—1647)，意大利数学家.

17 世纪上半叶微积分得到极大的发展. 关于积分运算(定面积体积，找重心)，卡瓦列里和罗贝瓦尔只需提一下，因为费马自己的研究走得相当远，他发现了经典积分过程. 另一方面，为了战胜时而剧烈的牙痛而解决轨迹线问题，帕斯卡意外地发现了求三角函数高次幂积分的过程. 我们所谈论的数学家的名字也能在微分运算中找到(相切问题). 通过他的"极大极小"方法，费马引进了无穷小数的概念. 拉格朗日和拉普拉斯是无穷小运算的真正缔造者，埃米·皮卡相信帕斯卡关于轨迹线的工作是无穷小积分运算的开端. 最初，莱布尼兹将他关于无穷小运算的公式潦草地写在费马某部手稿的复本上，就如他自己所说的，这个复本让他豁然开朗.

罗贝瓦尔(Roberval, Gilles Personne de, 1602—1675)，法国数学家、物理学家.

为了完整地说明帕斯卡的贡献，我们还得提及他在 28 岁时建造的第一台能做加减法运算的计算机器. 由于《论液体平衡》(*Traite de l' equilibre des*





liqueurs)这部著作,帕斯卡可与阿基米德同被誉为流体静力学的缔造者.也就不奇怪,他用于检测现在称做帕斯卡原理的桶,在立于皇家港口修道院墓地的小礼拜堂中得到展览.最后让我提一下关于大气压的试验,据推测,这是在梅森的影响下做的.梅森是一小群哲学家、数学家和物理学家的灵魂,在1666年科学院成立之前,这个团体代表了最早的小型但活跃的科学院.

那是同一个人能在哲学、数学和物理学都取得成就的幸运年代.哲学家马勒布朗士也拥有这样不寻常的直觉,他认为色彩与组成光的不同频率的振动数相关.

2 克莱洛的贡献

统治17世纪下半叶和18世纪初的数学家是荷兰的惠更斯,英国的牛顿和德国的莱布尼兹.牛顿和莱布尼兹因发现系统化无穷小微积分而闻名,而惠更斯则因在微分几何、有限应用力学,特别是关于光学(他最早提出和发展了与牛顿的粒子说相反的波动理论)方面的工作而著名.这段时期,牛顿证明星星和地球上的物体受同样的力学定律支配,即万有引力定律.它解释了行星、月球、彗星的运动和地球重力的存在性、潮汐的涨落等.由此引发了一场空前的科学革命.是牛顿的天才创造了一门全新的科学——天体力学.虽然这门科学最早诞生于英国,但是法国给了它特别肥沃的土壤以保证其将来的发展.为此,只需回忆一下对这门科学的发展做出最大贡献的科学家的名字:克莱洛、达朗贝尔、欧拉、拉格朗日、拉普拉斯、高斯、柯西、泊松、勒威耶、Tisserand,最后特别是亨利·庞加莱.

惠更斯(Huygens Christiaan,1629—1695),荷兰数学家、物理学家、天文学家.

牛顿(Newton,Isaac,1643—1727),英国数学家、物理学家、天文学家.

我要停下来介绍他们中的第一位克莱洛.克莱洛在21个兄弟姐妹中排行第二,父亲是数学老师.克莱洛展现了类似于帕斯卡的天赋;然而和帕斯卡不一样,他的早期工作并没有展示出后期工作的巨大意义.他在13岁之前就寄出了第一份通讯报告,在16岁就发表了一篇关于有双曲率直线(lines with double curvatures)的文章.在他18岁时,国王破格授命他为科学院力学部的成员.我不打算笼统地告诉大家他的纯数学研究工作,特别是关于解微分方程的部分(这些对研究微分方程的人应该都是熟知的),相反是关注那些使他闻名于世的工作.牛顿和惠更斯不但从理论上证明地球不是一个标准的而是两极稍扁的球面,而且给出了度量扁平度的方法.然而,当卡西尼1701年在比利牛斯山定出了巴黎子午线的弧度,他们的结论受到了质疑.经过混乱但总是生动的辩论,1736年科学院决定启动一项由莫佩蒂带领,意在确定拉普兰子午线弧度的探险.工作的条件极其艰难,还因大雪和极光更加恶化,研究团队(克莱洛也在其

克莱洛(Clairaut,Alexis Claude,1713—1765),法国科学家.

勒威耶(Le Verrier,Urbain Tean Juseph,1811—1877),法国天文学家.

中)得出的数值比卡西尼在法国得到的结果大得多,因而毫无疑问地证实了地球是两极稍扁的球体.自然的,莫佩蒂通过这次成功的探险获得了荣誉,他在一张肖像中头顶熊皮帽摆出一个用手按着地球仪的姿势.但是克莱洛继续思考地球两极稍扁的可能原因,尽可能从理论上确定受牛顿引力吸引的流体状行星可能有的形状.他的研究结果于1743年在《地球形状理论》(*La Theorie de la Figure de la Terre*)中出版,达朗贝尔称这本书是当时所做的所有事情的经典叙述,这是天体力学历史上值得记忆的日子.此外,克莱洛还解释了月球的运动并由此发展了牛顿的月球理论.他将这个领域中的结果概括进《月球理论》(*Theorie de la Lune*)一书并于1732年出版,两年后,他在书中加入数据表,就像方丹所说的,使它可能弄清天上月球运行的每一步.几年后,克莱洛因预测了哈雷彗星的下一次到来的日期而赢得了大众的认可和声誉.他解释由于土星引起的摄动将使哈雷彗星滞后大约100天,木星的影响使滞后再增加580天,预计彗星下一次掠过近日点大约在1759年4月13日,但请注意,由于许多其他不得不忽略的因素,预测的日期比实际日期晚了近一个月——哈雷彗星实际掠过近日点是在1759年3月13日.大约一个世纪之后,法国天文学家勒威耶通过定位一颗当时还不知道的行星(它是造成天王星扰动的主要原因)获得几乎同样的声誉.

莫佩蒂(Maupertuis, Dierre Louis Moreau de, 1698—1759),法国数学家、生物学家、物理学家.

方丹(Fontaine, Alexis, 1704—1771),法国数学家.

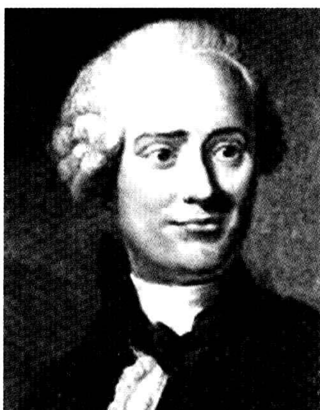
3 拉格朗日与达朗贝尔

18世纪的数学受欧拉和拉格朗日统治,达朗贝尔次之.

"数学王子"莱昂纳德·欧拉出生于巴塞尔,曾经旅居过圣彼得堡和柏林.他的天赋遍及数学的所有领域,他的工作有着重大而持久的影响.我将永远记得在阅读他的《无穷小分析引论》时所经历的愉悦,这本书是我中学最后一年获奖所得的礼物:它为我打开了一个全新的世界,为我更好地理解索尔邦和高等师范学校的课程作了准备.

达朗贝尔在许多不同的数学分支中都留下了他的足迹.一个以他的名字命名的著名代数定理断言有理方程的所有解(包括重根复解)的个数等于变量的最高次幂.虽然达朗贝尔关于这个结果的证明是错误的,但欧拉基于完全不同的方法的证明也并非毫无瑕疵.著名数学家高斯登上舞台后,给出了第一个正确的证明,后来柯西的出现则给这个定理带来一个既正确又非常简单的证明.在分析中我将只提及弦振动偏微分方程的第一个正确的表述,它是来自达朗贝尔的.最后,非常值得一提的是,在力学中达朗贝尔提出的一个原理(现在称做达朗贝尔原理)为拉格朗日的分析力学铺平了道路.





达朗贝尔



拉格朗日

拉格朗日出生于都灵的一个法国家庭，虽然他像欧拉一样，在柏林住了一些年，但他于 1787 年在巴黎永久定居，使法国人民将他视为自己最著名的人物.拉格朗日是那种在任何时代都真正重要的数学家之一.他的工作遍及数学的任何领域：在数论中，他证明了费马大定理在幂次为 4 的情形成立；在代数中，通过发展解多项式方程的有限递降法，即将多项式化约为一个次数更低的多项式，他为阿贝尔、高斯、伽罗华扫清了道路；另外，他还证明了五次多项式方程不能以求解三次、四次多项式方程的方法求解；在分析中，他给出了求解一阶偏微分方程的方法，同时提出奇异解的概念；在函数论中，他尝试了为无穷小运算建立一个严格的基础，但不是很成功，这个领域的基本原理还未像所期望的那样确切地发展，但其推论却被认为是正确的.然而，尽管没有完全成功，他的不考虑其几何或力学意义，以抽象的方式思考函数的方法，在现代函数论发展中仍有巨大的影响.拉格朗日归纳的天赋在他关于变分法的工作中显露无遗.

变分法是 18 世纪通过来自瑞士的两位数学家伯努利和欧拉的工作发展起来的.它的根源来自几何和力学中的一些问题，最简单的一个是求过同一张曲面上两点的最短路径.这里，未知的量不是一个数，而是更复杂的、一条由无限个点组成的线.莫佩蒂通过他的最小作用原理，将求解给定力场中粒子的运行轨迹的问题化约成极大极小问题，他给了这种类型的运算特殊的重要性.但不该忘记的是，费马那时已经将光学定律化约成相似的原理，根据这项原理，光线所选择的运行路径在时间上是最短的.通过在一未知的路径上作无穷小变分且给出这个变分的具体计算方法，拉格朗日将一个一般的方法发展为理论，因为解决其中每一问题都需要一个特殊的程序.

我将略过拉格朗日在天体力学中的工作，而花更多时间介绍他最重要的著

作《分析力学》(*Mecanique Analytique*)(1788).伽利略、笛卡儿、惠更斯、莱布尼兹、牛顿和达朗贝尔逐渐发展了现代力学的主要原理.但是确定一个受给定力支配的系统的运动轨迹,通常因考虑到不同力之间的未知关系而变得很复杂.在无摩擦的情形下,拉格朗日以其机灵的直觉完全克服了困难,给出一个确定所求解问题的运动轨迹方程的一般程序.为此,只需确定那个系统的主动力和系统在无穷小位移下该力所做的功.除了应用价值,这项奇妙的创造还有其显著的哲学意义,因为从力学特性的观点看,它完全阐明了质点系统的所有要义.在这个意义上,拉格朗日的天才可与解析几何创立者笛卡儿相媲美.

分析力学中所谓的拉格朗日方程为物理理论的各种力学解释提供了一个解析模型.从这个观点来看这项工作具有重大的哲学意义;然而,尽管这是19世纪最重要的工作,但它给人一种任何东西都可以由力学原理来解释的印象,和笛卡儿的任何东西都可以由几何语言来解释的信念一样的错误,这也就是现在它已被完全抛弃的原因.但是,这仍说明了数学可以给物理学家提供他们实现理论所需要的工具.

有些非凡的科学家认为,制造为无限序列现象提供洞见的结构(像拉格朗日所创造的结构)将会带来危险;他们害怕这种结构可能引起人们脱离现实.比如,伟大的几何学家彭赛列,因在力学中的工作而闻名.为避免使用拉格朗日方法,相反,他倾向于一步一步地,根据各种力之间相互作用和影响的所有细节,去分析它们的实际机理.同样类型的怀疑主义使得彭赛列拒绝使用解析几何,为此,他不得不使用经典几何原理去直接检查各种几何图形之间的关系.在接受最新的成果方面,的确有两类不同的头脑,两者对科学的发展同样重要,他们也并存于法国数学家中.

4 法国在数学中的优越性

法国在数学中的优越性早在18世纪末期就已显现,在法国大革命期间和19世纪初已变得非常清楚.这时期的伟大人物包括蒙日、拉普拉斯和勒让德.

拉普拉斯因在天体力学中的研究而闻名,其研究概括在他迷人的论文"宇宙系统之阐述"(*Exposition du Sysetme du Monde*)中.他阐明了下述奇特的观点:几乎所有天体现象甚至最精致的细节都隶属于科学决定论,据此,为了能够确定某一时刻宇宙质点的位置和速度,只需知道它在其他任何时刻的位置和速度,除此之外,假定我们还知道是什么原理控制着这些质点所受的力——以牛顿的万有引力为模型.数学物理的发展很长一段时间都受制于这个观点,直到最近,电磁学和原子物理才得以证明它是错误的,但这个结论仍旧对科学发展



有很大的影响.《概率的解析理论》(*Theorie Analytique des Probabilities*)(1812)是拉普拉斯又一篇非常重要的论文;这部著作中最重要的部分是处理概率的观念在最小方根理论中的应用,其概率已由勒让德提出.在研究椭圆体的弯曲中,拉普拉斯引进球面函数,通过这种方式我们可以表达任一定义在球面上的函数.我们不该忘记为牛顿势函数所满足的拉普拉斯的著名方程,这个方程在分析、几何、力学和物理的许多问题中都有不同寻常的重要性.

勒让德继欧拉之后为数论的复兴作出了重要贡献.虽然欧拉第一个发表了算术中的二次互反律,但勒让德作了清晰解释并部分地证明了它;因此这项定律以勒让德命名.高斯是发现同一定律的第三位数学家,也是第一位构造正确、证明完备的数学家.勒让德多年的重要工作,两卷本的小册子《关于椭圆积分》(*Sur les Integrales Elliptique*),于 1825 和 1826 年出版.书中他做了关于四次多项式平方根积分的完全研究,发展了表达它们的不同形式.这项工作使勒让德成为非凡的椭圆函数论的缔造者,同时也让雅可比和阿贝尔享有创立者的荣誉.最后,让我们提一下他的《几何基础》(*Lements Geometrie*)(1794),一部发行过许多版的著作迅速取代了欧几里得的理论;在非欧几何的历史中,这部著作有着绝对的重要性.



勒让德



蒙日

蒙日(Monge, Gaspurd, 1746—1818),法国数学家、化学家、机械理论学家.

蒙日是法国最好的几何学家之一.首先,通过创建现代射影几何,他加入到发展透视的悠久历史中去.透视的原理已为意大利文艺复兴时期的画家们所知,迪沙格和帕斯卡用它来研究圆锥曲线,随后法国几何学家 de la Hire(1640—1718)将它推广到圆的极点和极线理论(the theory of poles and polars of a circle).勒让德将射影几何系统化,通过非平坦曲面上的构造丰富了它.其次,蒙日的论文“分析在几何上的应用”(*Applieations de l' Analyse á la Geometrie*)极大地促进了

梅斯尼埃(Meusnier de la place, 1754—1793),法国数学家、物理学家、工程师.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

微分几何的发展,这个领域由于欧拉和梅斯尼埃关于曲面性质的工作而从笛卡儿的解析几何中分离出来;感谢蒙日给了我们曲率度量的概念和它在体积测定中的应用;这是他的想法:将一族数量非常大的曲面看做是单个偏微分方程的解的集合.他能够对极小曲面方程进行积分,极小曲面已经是且仍然是一个重要的研究课题,最早在普拉托实验中得到.蒙日在高等师范学校(这所学校作为女修道院于1795年创建)讲课时教授了他的理论,也在综合工科学校(拉格朗日和拉普拉斯也在那讲课)任教授.我很高兴有机会提到迪潘,因为他是和蒙日合作过的许多学生之一;迪潘因他的工作——《几何的发展》而出名,其中在曲面上一点处引进了公轭相切和指标的概念;同样,迪潘可被视为几何新分支的创造者.

普拉托(Plateau, Joseph Antoine Ferdinand, 1801—1883),比利时数学家、物理学家.

迪潘(Dupin, Pierre-Charles-Francois, 1784—1873),法国数学家、经济学家、教育家.

5 开创新方向

18世纪上半叶法国最知名的数学家有傅里叶、柯西、彭赛列和伽罗华.虽然各自的差别相当大,但他们都为科学开创了新方向.

傅里叶是数学物理方面公认的奠基人,我将略去他在代数方面的重要结果而直接给你们谈《热的数学理论》(*Theorie Mathematique de la Chaleur*),这部著作直到1822年才发表,但主要思想至少在1807年就已经形成了.伴随着这部著作,傅里叶开辟了数学分析中的一个新领域."不为古代几何学家所知,首次被笛卡儿用来描述曲线和曲面,"傅里叶说,"解析方程决不会限制在研究这些普遍现象,既然数学分析确定最多样的关系,度量时间、空间、力和温度,可以很肯定地说,它和自然界一样的广阔和丰富.在确证宇宙的同一性、简单性、稳定性中,它总是沿着相同的路径并给出相同的解释."不该忘记的是,对傅里叶来说,所有数学发现最丰富的源泉是大自然本身.举例来说,热的数学理论对纯数学的发展有着显著的影响,我们可以说傅里叶的观点是正确的.为了统一经常碰到的偏微分方程,傅里叶创立了三角级数理论.三角级数论引发了数量惊人的论文,旨在尝试为这个理论建立严格的数学基础同时完善和发展它.需要解决的基本问题是确定什么样的函数能够表示成傅里叶级数的形式.因为即使许多傅里叶自己的例子就很奇特,没多久数学家就感到真正迷惑了——一种类似于音乐家的迷惑,他们因发现,通过组合有限或无限多纯粹声音和它们的各种和谐音,能创造出任何不连续的声音序列而感到疑惑.这些不同寻常的结果将迫使数学家不得不再一次检查和提炼函数的概念,同时,一点一点地开始思考他们自己的科学的基础问题.这些工作带来一些不可思议的后果,虽然其中的内容还未完全显现.群论——一个经常令数学家挫败,引发许多恐怕仍然没解决

傅里叶(Fouries, Jean Baptiste Joseph, 1768—1830),法国数学家、物理学家.





的悖论的领域——是最终从这些努力中发展起来的数学分支之一;另一个源自同样努力的分支是单实变函数论,由 19 世纪末 20 世纪初法国数学家开创.

奥古斯丁·柯西,一位异常多产的理论家,在数学的数论、几何、分析和天体力学等领域都取得了成功.不同于欧拉,他先要弄清楚级数是否有意义,即级数是否收敛,然后才去探究级数本身.可以说,通过这种方式,他开辟了精确性的时代.柯西发现了怎样去确定变量的值以使级数收敛的一般规律(之后被阿达玛独立发现).复变函数理论的创立是柯西另一项伟大成就之一.因为三个世纪以来,虚数在数学界声名狼藉.它们最早出现在所有根都是实数的相悖情形时的三次方程的求根公式中,由意大利代数学家发现.然而,一旦研究者熟悉这些新的量,学会如何使用它们,那么就很容易得到关于实数的重要结果,而且有些结果用任何其他方法都无法得到.在接近 18 世纪尾声,瑞士数学家阿尔冈通过发现虚数的重要意义——为表达既包含长度又包含方向的平面向量提供可能性,从而解释了虚数的秘密.当柯西开始用一个复数取代两个实坐标来表示平面中的点时,他得到了复变函数的想法,复变函数将平面中的一点映射到平面中的另一点.通过这种方式柯西创造了一个全新的世界.这个世界中的元素被完美地组织在一起:就像 Cuvier(1769—1832,法国自然学家)能从史前动物的一块骨骼建构动物本身一样,数学家只要获知柯西函数在任意弧线上的值,不论弧线多么短,就能将函数重新建构出来.这个世界的完美秩序,非凡无比的和谐,特别是在数论中,得出一系列决定函数性质的定理及它们的许多应用,都给人留下无比深刻的印象.

阿尔冈(Argand, Tean Robert, 1768—1822),瑞士数学家.

柯西为更多的发现创造了合适的土壤,胜过他所参与的研究,衡量他贡献的指标是关于单复变函数论的一系列成果的数量.这系列中的一个定理,因为简洁而漂亮,几乎足以使刘维尔的名字流芳百世.这个主题的另一个定理,以埃米·皮卡命名,打开了那时还潜在的宏大视野,创造了一系列至今还长盛不衰的文章.

用一种不同于柯西的思路,德国数学家魏尔斯特拉斯也发展了一套单复变函数理论.很长一段时间以来,大家都认为他们各自的观点是不相关的,但是埃米·波莱尔在他一个最迷人的结果中证明这是不对的,柯西的观点更深刻地洞穿到问题的本质.波莱尔的确除去平面上一部分,如此一来以至平面上任何小的圆都不能保持完整,而在剩余的部分的内部他能构造一个函数,虽然满足所有柯西的要求,但是不满足魏尔斯特拉斯的定义.这样的单复变函数的存在性是以平面一个完整部分的存在性为条件的.通过他关于函数论的著名文集——这部文集现在和过去的贡献中包括各个国家的数学家——波莱尔自己对单复变函数论做出了许多贡献.


柯西


彭赛列

彭赛列(Poncelet, Jean-Victor, 1788—1867),法国数学家、力学家.

热尔岗(Gergonne, Joseph-Diea, 1771—1859),法国数学家.

施泰纳(Steiner, Takob, 1796—1863),瑞士数学家.

施陶特(Staudt, Karl Georg Christian Von, 1798—1867),德国数学家.

沙勒(Chasles, michel, 1793—1880),法国几何学家、数学史家.

伽罗华(Galois, Évariste, 1811—1832),法国数学家.

伴随着彭赛列我们进入了纯几何的时代.彭赛列被认为是射影几何的创立者.该领域的主题是研究物体在投影变换下保持不变的性质.他发现了既新颖又很实用的互反极变换,这个变换使得在满足下述奇特条件下可以将一个平面图形变换成另一个:新图形的边对应于原来图形的顶点,反过来也成立.经常地,这种类型的变换可将某图形所需探索性质化约到另一个更易探索性质的图形上.稍晚些,热尔岗用它来推导射影几何中非常重要的对偶原理.最后,彭赛列发现了连续原理,根据连续原理,如果一个图形有某种特性,假如考虑到不同元素之间的比例,即使在形变之下它仍将保持这种特性.通过许多简单的例子,柯西证明,彭赛列所阐述的这个原理是错误的;然而,如果以一种稍微不同的、更精确的方式来表述,这个原理事实上是正确的.因为很有帮助,这个原理经常被使用.在几何中,彭赛列的影响是非凡的:在德国,施泰纳和施陶特将他们的工作归因于彭赛列;在法国,沙勒(索尔邦高等几何学部第一位成员)是现代纯粹几何的杰出代表.对于沙勒,我们感谢他的重要的历史纪念性著作《几何发展的历史概述》(*L' Apercu historique sur le Develop pement de la Geometrie*),这本书为我们改正了不少错误观念.

在我们结束彭赛列的讨论之前,我注意到他在发展应用力学方面扮演了重要的角色,这门课他教了很长时间,先在梅斯(Metz),后在索尔邦.

伽罗华是科学历史上最不寻常的人物之一.两次在高等工科学校入学考试失败后,1831 年他被师范学校录取,但一年后就离开了.积极参与政治使他坐了几个月的牢;还未满 21 周岁,他便在一次由无意义的争吵所引发的决斗中死去.他曾将自己关于方程理论的数学发现发布在科学院的两份不同的通讯报告上,但后来两者都丢失了;幸运的是,他也将它们以小论文的形式发表在 1830 年的 *Bulletin de Ferussac* 上,还在去世前不久的一封信中和他的好友谈论过它们.



在他的论文中发现的一些其他结果,于 1846 年发表在刘维尔主编的杂志上.

他的工作的意义可以很快地做出解释. 16 世纪意大利代数学家塔尔塔利亚、卡尔达诺和费拉里用二次方根和三次方根解三次四次方程;然而,用同样的方式解更高次方程的根的努力都是徒劳的. 通过有些类型的方程的确能以同样的方式求解,拉格朗日、阿贝尔和高斯都对这个问题做出了伟大的贡献. 1826 年,阿贝尔首先表明,一般五次方程不能用根式求解. 通过这种方式,就变得很清楚,自从 16 世纪以来数学家就在努力解决的问题并没有得到正确的表述. 解决这个问题的荣誉属于伽罗华,他证明每一个方程都决定一定数目的根的置换,这些置换组成一个所谓的群;虽然作用于根,但这些置换不会破坏它们之间的有理相互作用("有理相互作用"这个概念需作附加解释). 这个群的特征决定了方程的基本性质,不论是否可能找到根;在一般情况下,也决定了辅助方程的特征,而求解它将会使原方程得到解决. 从他自己的观点出发,伽罗华轻而易举地得到他的前辈们的结果,且成功地将它们整合到他自己的理论中去.

塔尔塔利亚(Tartaglia, Nicolò, 1499—1557),意大利数学家、军事科学家.
卡尔达诺(Cardano, Girolamo, 1501—1576),意大利学者.
费拉里(Ferrari, Ludovico, 1522—1565),意大利数学家.

置换群理论,即一定数目对象的置换所组成的群,由柯西创立,通过伽罗华的工作才展示了它的全部价值. 伽罗华完善了它的重要方面,解释了普通群的基本意义. 更进一步,他通过引进新的虚数(伽罗华虚数)丰富了数论,每一个虚数都联系到一个素数的次幂;伽罗华的名字不仅在方程理论里会经常碰到,而且在现代代数中也是如此. 他寄给朋友 Chevalier 的信件使我们清楚他在分析中有着他在代数中同样多的重要成果,且他在阿贝尔积分方面的工作比著名德国数学家黎曼在同一方面的工作要早 25 年. 虽然伽罗华的早逝使科学界蒙受巨大损失而使我感到悲伤,但我还是得说,就像埃米·皮卡曾经提过的"当遭遇如此短暂而动荡的一生,我们对这颗在科学中留下如此深刻足迹的非凡头脑更加充满敬意".

因为伽罗华理论才使我们能够解释,允许虚数出现在有实根的三次方程求根公式中的奇迹;事实上,我们可以证明,如果一个方程的根都是实数,且方程能用根式求解,那它的解中只含有二次方根. 用同样的方法,我们可以表明一些古老的问题——比如倍方问题和三等分角问题——不能由尺规作图的方法得到. 若尔当以他重要的著作《替换论》(*Traite des Substitutions*)为伽罗华立了一座纪念碑.

韦西奥(Vessiot, Ernest, 1865—1952),法国数学家.

因为既简单又深刻,伽罗华的主要思想在代数方程之外的其他地方也得到了应用. 比如,埃米·皮卡和韦西奥认为它在线性方程的积分中非常重要. 值得一提的是,德拉什和韦西奥尝试将伽罗华理论推广到最一般的微分方程,但是遇到了困难,只有改变原先的理论或者至少要牺牲它一部分惊人的简洁性才能克服其中的困难.

伽罗华之后科学的发展表明群论在数学和物理各种不同分支中的重要性

在不断地增长.挪威数学家索弗斯·李,变换群理论的创始人,将它们引进几何和分析中,作为伽罗华的一个崇拜者,他将关于变换群的纪念著作于1889年献给高等师范学校.事实上,关于伽罗华理论的发展、精炼、推广和寻找新应用的最重要成果都在法国产生.庞加莱宣称群的观念已经存在于几何的精神中;如果两个几何图形中任一个都与第三个相互等价则称它们等价,这个公理事实上等价于存在一个规范几何的群,更准确地说,经过一系列过程一个图形就可以转变成它的等价图形.群论能给我们用以表达"等价图形"的所有具体、相关的含意,这有着不寻常的重要性;就像德国伟大数学家菲利克斯·克莱因1872年所提出的,这确实表明存在无限多的几何,每种几何都由一个群规范,同时,每一种几何都可以独立地考察研究,不必求助于初等几何.这个框架包含了射影几何,其中两个图形被认为是等价的,如果其中一个可以通过一系列的射影变换得到另一个.

6 光辉灿烂的纪念碑

安培(Ampère, André-Marie, 1775—1836),法国物理学家、数学家、化学家.

拉梅(Lamé, Gabriel, 1795—1870),法国数学家.

泊松(Poisson, Siméon-Denis, 1781—1840),法国数学家、力学家、物理学家.

弗雷内尔(Fresnel, Augustin Jean, 1788—1827),法国数学家、物理学家、工程师.

伽罗华逝世已有100年了.在这期间数学得到了极大的发展,大量的论文发表了,但不得不说,有些文章在图书馆中占据着不该有的位置.有些理论,在伽罗华的时代就发现,之后得到深刻的探究,它们中有些还渗透到其他数学领域中.总而言之,数学像其他科学一样,永远在变化,极富戏剧性,因此,要对数学的现状有真正的洞察,无论对谁而言,都变得很困难.越来越少的头脑能够既在纯数学又在应用数学中做出重大发现.出现像法国的安培这样既是物理学家又是著名数学家的天才是很罕见的.在物理学中,他创立了电动力学;在数学中,他和蒙日分享了创造二阶偏微分方程理论的荣誉.法国的拉梅是分析学家、几何学家和弹性理论的奠基人;而法国的泊松因分析学和数学物理方面的工作而著名.弗雷内尔,物理光学的创造者——他的工作最终,至少直到量子物理出现,保证光的模理论的胜利——可同时看做是数学家.

与其给你们一列冗长乏味的名单,还不如让我们把注意力投向几位法国当代最伟大的数学家,他们是我的老师.有机会表达对他们的敬意我感到很荣幸.

查尔斯·埃尔米特在进入高等工科学校不久,就给著名教授雅可比写信,同时寄给他一篇关于分类阿贝尔超越函数的论文,这些函数和最一般代数微分的积分相关.雅可比和阿贝尔一起是椭圆函数论的创立者.他曾经在类似的情况下,得到勒让德的友好回复,对年轻的埃尔米特了不起的工作表示了祝贺.这是两位伟大数学家之间有规律的通信的开端.埃尔米特在24岁时寄给雅可比关于高等代数的发现,最终确保他跻身于最著名几何学家之列.建立在最著名的高斯的结果之上,他自信地发展最一般形式下的形状的代数理论(Algebraic



Theory of Shapes)，且在以非连续为特征的数论中引进连续变量．他引进带不定共轭项的二次型(现在叫做埃尔米特型)，这就是为什么他的名字在量子物理著作中经常碰到的原因．1873年，埃尔米特因发现 e 的超越性而闻名于世，它是Neper算法的基础(超越数，即不满足任何系数为有理数的代数方程的数，其存在性最早由刘维尔证明)．因为埃尔米特的结果给人印象深刻，有人期望他去证明 π 的超越性，由此完全断绝由尺规作图可以平方化圆的希望；但是，德国数学家林德曼从埃尔米特的方法中得到灵感，以适当的方式改进了这个方法，并得到一个证明，为自己争得了荣誉．

林德曼(Lindemann，Carl Louis Ferdinand，1852—1939)，德国数学家．



安培



埃尔米特

埃尔米特总是给听众留下深刻的印象．“没有人会忘记埃尔米特讲演中布道式的语调，”著名数学家班勒卫说道，“以及当我们聆听埃尔米特谈论一项不可思议的发现或一些亟待去探索的东西都会经历到的对美的感受和启示．他的话总能极大地开阔科学视野，传达了对完美理想的爱和敬意．”“每当我有机会聆听埃尔米特的教诲，眼前总会浮现出由数学沉思带来的平静而纯粹的欢乐，正如贝多芬内心感受自己的音乐时必定会体验到的那样．”

达布(Darboux，Jean-Gaston，1842—1917)，法国数学家．

达布既是分析学家又是几何学家．虽然他是分析学中一些结果的开创者，但是我将不讨论他这方面的工作，因为在几何方面的工作使他出了名．他当然不是那种通过推崇分析来避免使几何美失色的几何学家，也不是倾向于将几何简化为计算，而不顾几何意义或对几何意义不感兴趣的分析学家．在这个意义上，他跟随蒙日的足迹，将精美、充分发展的几何直觉和分析的纯熟应用结合起来．他所有的方法都极其优雅，非常适合所考察的主题．在索尔邦高等几何部教书时(他继任了沙勒的位置)，他经常地且充满敬意地谈论三维正交系统，高兴地强调拉梅的工作的重要性．达布同样经常地谈到平面形变的理论，这个理论最早起源于高斯的《关于曲面的研究》(*Disqusitiones circa Superficies Curvas*)，即使

博内(Bonnet, Pierre-Ossian, 1819—1892),法国数学家.

在达布之前,这也是一个法国数学家们做出重要工作的领域,其中博内当然是值得一提的.最后,达布说明单位坐标系统的威力,即与所考察图形相关的坐标系,由于有了群论,埃利·嘉当进一步发扬了这个方法,使它适用于为广义相对论发展的需要而创立的最广泛的空间.达布在几何发展中有着巨大的影响力,在他无数学生和追随者中,我只打算提一下罗马尼亚著名几何学家 Tzitzeica,他是《巴尔干数学评论》的创始人之一,他于最近去世,却仍然为整个科学界深切悼念.达布在这个领域的经典著作《曲面理论》(*Theorie des Surfaces*)是为表达对几何和分析的敬意而树立的光辉灿烂的纪念碑.

有一个故事讲的是,当一个年轻的德国数学家因为拉格朗日拒绝认为高斯是德国最伟大的几何学家而感到疑惑时,拉格朗日告诉他:"不,他不可能是德国最伟大的几何学家,因为他是欧洲最伟大的几何学家!"以同样的方式,我们可以说亨利·庞加莱不仅是个伟大的数学家,而且就是数学本身.我们不可能找到任何一个数学分支,甚至任何一个物理分支是他没有留下任何足迹,或没有复兴过,或没有推演出一个全新的领域的.在创造富克斯函数之后,他使用一致函数以同样的参数表达代数曲面上点的坐标,由此得到之前只适用于特殊类型曲面的结果.他在那个时候用这样的方式解决一致性问题是很勇敢的.他是多复变函数论的先驱,创立了实域中的微分方程理论;因为这个理论,他可以重构天体力学理论的方法,来研究这个领域中问题的周期解和考察稳定性问题.在代数拓扑学中,这是几何中研究在连续变换下保持不变的那部分性质的分支,庞加莱署名的几篇论文,成为几乎所有后来理论的出发点.在索尔邦大学,他关于数学物理的各个方面的讲座,影响了后来由麦克尔逊实验引发的想法.因为他的早逝,科学界失去了一位最著名的领袖.他的科学哲学著作《科学与假设》(*La Science et l' Hypothese*)和《科学的价值》(*La Valeur de la Science*),被翻译成许多种语言,闻名全世界.在某种意义上,庞加莱可与帕斯卡相媲美,这里引用庞加莱自己的话:"思想是漫漫长夜的瞬间一闪,但这一闪就意味着一切."要发展庞加莱的所有想法,要探索他丰富多彩的全部著作内容,需要花费很长的时间.

阿佩尔(Appell, Paul, 1855—1930),法国数学家、数学物理学家.

古尔萨(Goursat, É ouard-Jean-Baptiste, 1858—1936),法国数学家.

最后,我将提及阿佩尔和古尔萨,前者是《有理力学论》(*Traite de Mecanique Rationnelle*)的作者,后者是《微分和积分论》(*Traite de Calcul Differentiel et Integral*)的作者,还将再次提及埃米·皮卡,著名人物中还在世的一位.两年前,埃米·皮卡和德国伟大数学家大卫·希尔伯特同时获得了由瑞典米塔-列夫勒研究所颁发的金质奖章.就在几周前,在埃米·皮卡成为科学院成员 50 周年庆祝会上,埃米·波莱尔谈到了他的科学工作.我已经提到由他的名字命名的著名定理,和他发展伽罗华理论的那些著作.他的关于两个变量的代数函数的工作表述了代数几何的基础,这个几何分支在意大利得到极大的发展.从过去一个世纪



看意大利几何学家的观点确实比埃米·皮卡的观点更清晰;但是,就如埃米·波莱尔所提到的,如果没有埃米·皮卡的贡献,代数几何肯定会变得不完整.

7 法国数学的光荣

由埃尔米特、达布、庞加莱和皮卡所带来的法国数学的光荣,并没有变得黯然失色.事实上,法国数学的光辉和它曾经有过的贡献一样灿烂.因为时间简短,为说明这个事实我不得不限制只提少数几个名字.

柯尼希是位很好的几何学家,他某些优雅的工作可与达布的成果相媲美.通过构造新的超越数,班勒卫解决了一个连庞加莱都无法解决的问题;庞加莱将班勒卫在分析中的工作概括如下:"数学是块完好组织的大陆,它的国家得到统一;班勒卫的工作是海洋中一个壮丽的岛屿."但是这样的评价是不完全的,因为班勒卫——在很长一段时间里,在高等工科学校教授力学——同样显著地推进了力学的发展.除此之外,它的理论研究还促进航天学的发展,以致我们可以这样说,由于班勒卫,航天学是法国独有的创造.

雅克·阿达玛的工作既多样又重要:在算术中,他研究关于复杂的素数分布问题的黎曼函数;在几何中,他研究负曲率空间的测地线;在分析中,他出版了关于数学物理中偏微分方程的著作.同时,他也极大地刺激了变分法和泛函分析,这是由意大利数学家沃尔特拉创立的新科学.最后,他在法兰西学院的讨论班,吸引全世界数学家前来展示他们的最新研究成果,影响了数学中的国际合作.因为他还很年轻,可以肯定地说,他将要做的工作还远未结束.

法国复变函数的研究一直以来都很成功.这里我要提一下埃米·波莱尔,英年早逝的分析学家法图,因正则函数族的理论而闻名的蒙泰尔,因关于有理函数提升的工作而著名的朱利亚,等等.

实变函数理论几乎独自根源于法国.由若尔当的《分析学》(*Traite d' Analyse*)所设立(像埃米·皮卡的同名论著一样,有着深远的国际影响),由埃米·波莱尔的工作所创立,由亨利·勒贝格(他定义了集合的度量)、勒雷·贝尔(他引进现在以他名字命名的积分)以及当儒瓦(totalization 理论的创立者)所引进,使得它在一个被忽视很长一段时间的领域中引起了意想不到的和谐效果,见证了它的创造者们的勇气和天才.

我不得不提到弗雷歇的抽象空间理论、布利冈的无穷小几何、埃利·嘉当在

柯尼希(Koenigs, Gabriel, 1858—1931),法国数学家、物理学家.

沃尔特拉(Volterra, Vito, 1860—1940),意大利数学家.

法图(Fatou, Pierre Joseph Louis, 1878—1929),法国数学家.

蒙泰尔(Montel, Paul Antoine, 1876—1995),法国数学家.

朱利亚(Julia, Gaston Maurice, 1893—1978),法国数学家.

弗雷歇(Fréchet, Maurice-René, 1878—1973),法国数学家.

布利冈(Bouligand, Georges Louis, 1889—),法国数学家.

达尔穆瓦(Darmois, Georges, 1888—1960),法国数学家.

几何和分析中的工作,最后一位的工作不适合由我来作评价.

庞加莱学院来源于法国对数学物理领域研究新的狂热.埃米·波莱尔,概率论的灵魂人物,开始出版这个领域中一系列值得称赞的著作.就像函数论中那样,在这个系列中弗雷歇,保罗·列维和达尔穆瓦发表了他们杰出的研究成果.理论物理学部由 Louis de Broglie 主管,他是波动力学的创立者,重建了原子物理学,调和了光的波动性和粒子性.我不会忘记提及力学部,由维拉领导,他因流体力学中的工作而出名,他还是国际著名汇编《数学科学回忆录》(*Memorial des Sciences Mathematiques*)的编辑和《纯粹和应用数学期刊》(*Journal de Mathematiques Pures et Appliauees*)的主编,几乎有一个世纪,这个期刊最初由刘维尔创办,在相当长的时间里由若尔当做编辑.

维拉(Villat, Henri-René-Pierre, 1879—1972),法国数学家.

如果不提到综合工科学校和高等师范学校,法国数学活动的描述将是不完整的.一个世纪以来,伟大的法国数学家们都感谢这两所大学对他们的教育.最近半个世纪,这个重要的角色几乎为高等师范学校所垄断,即使在整整 50 年前李已经将它看做是法国数学的摇篮.来自其他国家的年轻天才来到这里和他们的法国同行享受着同样的教育.这就是不能不将我前面已经提到过的 Georges Tzitzeica 看做是法国数学家的原因.基于同样的理由,我倾向于将我的好朋友 Mihailo Petrovic 归入法国数学家,他是南斯拉夫数学的老前辈.他因为在算术、代数和分析中发明了谱方法的伟大原创性工作而得到公认,他还创立了一般现象学,这是一个系统研究解析模型(能用于同时表述几个明显不同的物理理论的模型)存在性问题的领域.我希望你们不会反对我将他的工作归到法国数学的伟大成就中.

由于高等师范学校,年轻数学家已经准备接替年老的数学家.你们可能会觉得要提名还为时过早,但他们中有些确实已经非常著名.我只想提一下雅克·埃尔布朗,他的工作因早逝而被无情地打断,但这并不妨碍宣布一个伟大数学家的诞生,这可能类似于伽罗华吧!

女士们先生们,是该到我结束演讲的时候了,我已经花去大家非常多的注意力.最后,我只想作一个很一般的总结.

有别于其他科学,数学通过一系列连续的抽象得到发展.为了避免错误,数学家被迫去发现孤立问题的本质和所思考的实体的关系.如果走向极端,这个过程就会如一个笑话所说的那样:数学家是既不知道他所谈论的又不知道他所



谈论的是否存在的科学家.但是,法国数学家不会使自己远离现实;他们确实知道,虽然逻辑是必要的,但无论如何不是关键.数学活动和人类的其他类型活动一样,需要找到一个有价值的平衡点:正确思维无疑很重要,但是提炼出正确的问题更重要.在这个意义上,我们可以自由地说:法国数学家不仅永远知道他们所谈论的问题,而且有着正确地选择最基本问题的直觉;这些问题的解决将在整个科学的发展中产生最强烈的影响.

九十九年后的庞加莱猜想①

第六章

2 维流形,或曲面的拓扑早在 19 世纪就已经清楚了.②事实上,可用一个简单的表列出所有可能的光滑紧可定向曲面.对任何一个这种曲面可以定义一个亏格 $g \geq 0$,取值于非负整数,直观上亏格可理解为洞的个数(图 6.1),两个这种曲面之间可以光滑的方式建立 1 - 1 对应的主要条件是它们的亏格相同.

高维的相应问题要困难得多.庞加莱或许是第一个试图对 3 维流形作类似研究的人.这种流形的最基本例子是 3 维单位球面,即,4 维欧氏空间中离原点距离恰为 1 的所有点的集合

$$x^2 + y^2 + z^2 + w^2 = 1$$

他注意到 2 维球面区别于其他曲面之处在于球面上的每条简单闭曲线都可以在球面上连续地形变收缩到一个点.1904 年,他提出了 3 维情形的相应问题.③用更现代的语言,可表述如下:

> 若一个光滑紧 3 维流形 M^3 具有如下性质:这个流形中的每条简单闭曲线可以连续地收缩为一个点,M^3 同胚于 S^3 吗?

---

① 原题:The Poincaré Conjecture 99 Years Later: A Progress Report.作者:John Milnor.译自:http://www.math.sunysb.edu/-jack/PREPR/poincaré03.pdf.

② 定义及相关材料,参见 Massey,或 Munkres 1975,以及 Thurston 1977(人名,或人名后有年代,均指本文末的参考文献.下同).——原注

③ 参见庞加莱,第 486,498 页,以及 370 页.——原注





图 6.1　亏格为 0,1,2 的曲面的简图

他以很高的预见性评论道:“但是,离我们解决这个问题还相当遥远.”自此以后,每个单连通闭 3 维流形同胚于 3 维球面这个假定便以庞加莱猜想而知名.它一直激励着拓扑学家们,试图证明这个猜想的许多努力也使我们得以对流形的拓扑有更多的了解.①

1　最初的失误

最初,庞加莱猜想字面上如此简单明了以至数学家们在此屡屡出错.提出这个问题的 4 年前,也就是 1900 年,庞加莱的以下错误定理使他自己成为在此猜想上失误的第一人.②

每个具有 n 维球面的同调群的紧多面体流形实际上同胚于 n 维球面.

但在作了进一步研究后,他发现了上述断言的一个漂亮的反例.几何上可以如下描述这个例子.考虑画在 2 维单位球面上的所有可能的正 20 面体.要确定这个族中的一个特定的 20 面体,需要 3 个参数.例如,确定位于球面上的一个顶点要 2 个参数,而另一个参数则用来确定指向相邻顶点的方向.因此,每个这种 20 面体是由所有这种 20 面体构成的 3 维流形中的一个“点”.③这个流形满足上述错误定理中的条件:用同调论方法无法将它与 3 维球面区分开来.然而他能证明它不是球面,用的方法是构造一条简单闭曲线,这条简单闭曲线在 M^3 中无法收缩为一个点.其构造并不困难:选取某个特定的 20 面体,考虑它围绕一个顶点作 $2\pi/5$ 逆时针旋转时得到的象.这样就在 M^3 中定义了一条不能收缩到一个点的简单闭曲线.

下一个重要的错误定理是由 J·H·C·怀特海在 1934 年给出的.作为其所谓庞加莱猜想证明的一部分,他断言每个可缩(即可连续收缩为一个点)开的 3 维流形同胚于欧氏空间.像庞加莱一样,他后来也发现了自己“定理”的一个反例,

J·H·C·怀特海 (Whitehead, John Henry Constantine, 1904—1960),英国数学家.

① 如果要了解历史上试图解决庞加莱猜想而作的各种尝试,请参见 Papakyriakopoulos, Birman, Jakobsche, Thickstun, Gillman 和 Rolfsen, Gabai 1995, Rourke, 以及 Poénaru.——原注

② 参见庞加莱,第 486,498 页,以及 370 页.——原注

③ 更准确地,这个 M^3 可定义为陪集空间 $SO(3)/I_{60}$,其中 $SO(3)$ 是 3 维欧氏空间所有旋转构成的群,而 I_{60} 则是把一个特定正 20 面体变到自身的 60 个旋转所构成的群.基本群 $\pi_1(M^3)$,定义为 M^3 中从一点到自身的所有闭道路的同伦类所构成的群,是一个阶为 120 的完满群(perfect group).——原注

从而大大增加了我们对流形拓扑的了解(参见 J·H·C·怀特海,21~50 页).他的反例可以简短地描述如下.从两个不相交的实心环体 T_0 和 T_1^* 开始着手,它们嵌入在 3 维球中,如图 6.2 所示,每个实心环体本身不打结,但是,它们却组成一个非平凡的环链,并且它们的链接数(linking number)为 0.因为 T_1^* 是不打结的,它的补空间 $T_1 = S^3 \setminus (T_1^*)$的内部是另一个包含 T_0 的不打结的实心环体.选取一个 3 维球面的同胚映射 h,它把 T_0 映到这个更大的实心环体 T_1 上.那么我们可以归纳地构造出 S^3 中不打结的实心环体的递增序列

$$T_0 \subset T_1 \subset T_2 \subset \cdots$$

其中 $T_{j+1} = h(T_j)$.这个递增序列的并 $M^3 = \bigcup T_j$ 就是所要的怀特海的反例.这个流形可缩但却不同胚于欧氏空间.要证明 $\pi_1(M^3) = 0$,注意到 T_0 中的每个闭道路可在更大的实心环体 T_1 中收缩为一个点(在收缩过程中,可能会出现自相交的闭道路).但 M^3 中的每个闭道路必定含于某个 T_j 中,因此可在 $T_{j+1} \subset M^3$ 中收缩为一个点.另一方面,M^3 不同胚于 3 维欧氏空间,因为若 $K \subset M^3$ 是任一包含 T_0 的紧子集,可以证明补集 $M^3 \setminus K$ 不是单连通的.

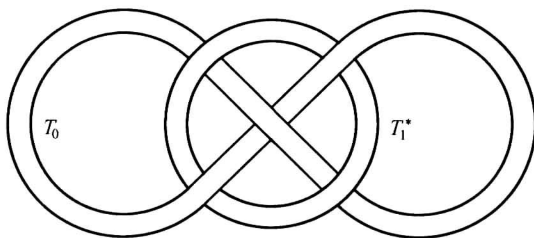


图 6.2　怀特海链环

自此以后,曾有许多关于庞加莱猜想的错误证明.其中的一些依赖于细微难辨的错误命题.可以看一看 Bing 的文章,它对 3 维拓扑中一些容易出错之处作了精彩的描述.

2　高维情形

在 20 世纪 50 年代末期和 60 年代初期出现了一批惊人的成果.人们发现实际上高维流形的研究比 3 维流形要容易.之所以如此,理由之一是:基本群对研究任意维流形都很重要,哪怕它是平凡群也是如此.基本群的生成元间的关系对应于映入流形的 2 维圆盘.在 5 维或更高维的情形,可把这些圆盘扰动到一般位置使其互不相交,也不会自相交;但在 3 或 4 维情形,相交可能是不可避免的,从而导致一些严重的困难.

1960 年 Stephen Smale 宣布了在高维情形庞加莱猜想的一个证明.紧随其后



的是 John Stallings,他使用了一种完全不同的方法;还有 Andrew Wallace,他沿着一条完全类似于 Smale 方法的路线进行研究.

让我先叙述 Stallings 的结果,它的假设较弱,证明较易,然而结论也较弱.他假定维数大于或等于 7;后来 Christopher Zeeman 把它的论证推广到了维数为 5 和 6 的情形.

Stallings-Zeeman 定理 如果 M^n 是维数 $n \geqslant 5$ 的有限单纯复形,它与球面 S^n 伦型相同,并且局部地逐段线性同胚于欧氏空间 $\mathbf{R}^n$,则 M^n 同胚于 S^n,其同胚映射除去一点外是逐段线性的.换句话说,补空间 $M^n \setminus$ (点)与 $\mathbf{R}^n$ 逐段线性同胚.

证明的方法是把所有的难点都推到某一点上,从而在此点附近无法控制同胚的性态.

Smale 的证明以及稍后 Wallace 给出的与前者紧密相关的证明,颇依赖于各种可微方法,即从一个 n 维球出发以逐步添加上环柄的方法来归纳地构造流形(图 6.3).这里把一个 k - 环柄安装到带边流形 M^n 的方法是:先用一个(k - 1)维的边缘球面到 M^n 的边缘中的粘接同胚,把一个 k 维胞腔粘接上去,然后再加厚并使拐角光滑以得到一个较大的带边流形.重新安排并消去这样的环柄就完成了证明.

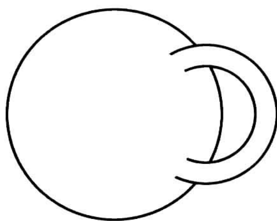


图 6.3 粘接了一个环柄的 3 维球体

Smale 定理 如果 M^n 是一个维数 $n \geqslant 5$ 的可微同伦球面,则 M^n 同胚于 S^n.事实上,M^n 微分同胚于这样一个流形,它是由两个闭 n 维球的边缘在适当的微分同胚下黏合而成的.

Wallace 也证明了这个定理,至少 $n \geqslant 6$ 时是这样的.(应当注意到,5 维的情形是特别困难的.)

对于困难得多的 4 维情形,人们等待了 20 年才等来 Michael Freedman 的工作.这里 Smale 和 Wallace 用过的可微方法和 Stallings 与 Zeeman 用过的逐段线性方法都完全没有用.Freedman 无所限制地使用了不可微方法,不仅证明了 4 维的庞加莱猜想而且给出了闭的单连通 4 维拓扑流形的一个完全分类.这样一个流形的整上同调群 H^2 是个自由阿贝尔群,Freedman 只需要两个不变量:杯积 $\beta: H^2 \otimes H^2 \rightarrow H^4 \cong Z$ 是个行列式为 ± 1 的对称双线性形式,同时 Kirby-Siebeni-

nann 不变量 k 是个 mod 2 整数,它为 0 的充要条件是乘积流形 $M^4 \times R$ 可被赋予微分结构.

Freedman 定理 两个闭的单连通 4 维流形同胚的充要条件是它们具有相同的双线性形式 β 和相同的 Kirby-Siebenmann 不变量 k.任何一个 β 均可由这样一个流形来实现.如果对某个 $x \in H^2 \beta(x \otimes x)$ 为奇数,则 k 可能取的两个值中任一个值也可实现.但是,如果 $\beta(x \otimes x)$ 总是偶数,则 k 由 β 决定,它与 β 的符号差的 1/8 同余.

特别,如果 M^4 是个同伦球面,则 $H^2 = 0, k = 0$,从而 M^4 同胚于 S^4.应该注意的是,逐段线性理论与微分拓扑理论在 4 维情形要困难得多.目前还不知道是否每个光滑同伦 4 维球面一定与 S^4 微分同胚,也不知道 $k = 0$ 时什么 4 维流形会有微分结构,也不知道这个结构何时是本质上唯一的.关于这些问题主要结果是由 Simon Donaldson 获得的.反映该问题复杂程度的一个例子是,Freedman 利用 Donaldson 的工作,证明了 $\mathbf{R}^4$ 具有不可数个不等价的微分结构.

在 3 维时,拓扑理论、逐段线性理论与微分拓扑理论之间的差异消失(参阅 Hirsch, Munkres 1960,及 Moise).但是基本群引起的困难却加重了.

3 Thurston 几何化纲领

在 2 维情形,每个光滑紧曲面具有一个漂亮的几何结构,这个几何结构在亏格为 0 时是正常曲率的球面,在亏格为 1 时是曲率恒为的平坦环面,而当亏格大于等于 2 时,是曲率为负常数的曲面.William Thurston 1983 年的一个影响广泛的猜想断言,类似的事情在 3 维也是对的.他的猜想断言,每个可定向的 3 维紧流形可以沿 2 维球面和环面切开,从而分解为一些本质上唯一的面.它们中的每一块都具有一个简单的几何结构.在 Thurston 的纲领中,有 8 种可能的 3 维几何.其中 6 种现已充分了解了,①而且在负常曲率几何方面也已取得很大的进展.②但是,对应于正常曲率的第 8 种几何却还少有研究.对于这种几何,我们有下述庞加莱猜想的推广:

Thurston 椭圆化猜想 每个具有有限基本群的闭 3 维流形有一正常曲率的度量,因而它同胚于商空间 S^3/Γ,其中 $\Gamma \subset SO(4)$为一个自由作用在 S^3 上的有限旋转群.

庞加莱猜想对应于特殊情形($\Gamma \cong \pi_1(M_3)$为平凡的情形).$SO(4)$的所有子群 Γ 在很久以前已由 Hopf 分类,但是这个猜想却远未解决.

---

① 例如,可见 Gordon 和 Heil, Auslander 和 Johnson, Scott, Tukia. Gabai 1992 以及 Casson 和 Jungreis.——原注

② 见 Sullivan, Morgan, Thurston 1986, McMullen,以及 Otal. Haken 和 Waldhausen 的前驱性论文是这方面工作的基础.——原注





4 微分几何方法和微分方程方法

给定一个光滑3维流形,它上面的所有黎曼度量构成一个无限维空间.近年来,通过研究这个空间的几何来研究几何化问题(从而研究庞加莱猜想)的尝试已有好几次.

根据定义,黎曼流形上一条道路 γ 的长度,是利用度量张量表成积分 $\int_\gamma ds = \int_\gamma \sqrt{\sum g_{ij} dx^i dx^j}$.从这个度量张量的1阶和2阶导数,可以算出Ricci曲率张量 R_{ij},以及标量张量 R.(例如,对平坦欧氏空间,$R_{ij} = R = 0$,而对半径为 r 的3维球面,Ricci曲率 $R_{ij} = 2g_{ij}/r^2$,标量曲率为 $R = 6/r^2$)

基于Hidehiko Yamabe的早期工作,Michael Anderson提出一种方法,将全标量曲率

$$S = \iiint_{M^3} R\mathrm{d}N$$

作为全体具有单位体积的光滑黎曼度量所成空间上的泛函来研究.这个泛函的临界点就是常曲率度量.

哈密尔顿提出另一种方法,研究Ricci流.即,下述微分方程的解

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

换句话说.度量被要求随时间变化,使得距离在正曲率方向减少.这本质上是一个抛物型微分方程,与物理学家研究的热方程很类似:假如我们在一个冷的杆子的一端加热,那么热会沿着杆子逐渐地流动直至达到均匀的温度.类似地,对基本群有限的3维流形.人们最初的期望是正曲率会沿着Ricci流逐渐扩散,达到极限(经过一个重整化后),流形应该会有常曲率.

如果从一个正Ricci曲率的3维流形出发,哈密尔顿可以实现上述设想,并构造一个常曲率的度量,因此解决了椭圆化猜想的一个很特殊的情形.一般情形,会遇到严重的困难,因为这个流会趋向于一些奇点.然而,他猜测这些奇点必定具有很特殊的形式,从而上述方法仍可用来构造常曲率度量.

3个月前,圣彼得堡的佩雷尔曼贴出一份预印本,描述了一种方法来解决哈密尔顿方案中一些主要难点,并提出解决全部椭圆化猜想的一种途径.专家对这个声明的最初反应是谨慎的乐观.不过,鉴于历史上这个领域中错误证明如此之多,在所有的细节得到仔细解释与验证之前,没有人敢肯定它是对的.

庞加莱猜想可能已被证明[①]

第七章

一位俄国数学家报道，他已经解决了庞加莱猜想——数学中最著名的未解决问题之一.

这位数学家，在圣彼得堡俄罗斯科学院斯捷克洛夫数学研究所的佩雷尔曼博士，将要在一系列文章中陆续描述他的工作.

他的工作被完全地验证则是几个月之后的事了.但如果是对的，它将证明一个困扰了数学家几乎一个世纪的问题，它的影响将遍及几何与物理.

如果他的文章被一份要审稿的研究杂志接受发表，并且经过两年时间(广大数学家的)广泛审查后仍存活，那么佩雷尔曼博士将获得100万美元奖金，这笔奖金是由位于美国麻省剑桥的 Clay 数学研究所赞助，为被该研究所确定为7个千禧年最重要的未解决的数学问题中的每一个而设立的.

当佩雷尔曼博士于2002年11月在一个因特网预印本服务器(Internet preprint server)上张贴了报告他的结果的系列文章的第一篇之后，关于他解决庞加莱猜想的消息就传开了.

上周，在麻省理工学院，他对满座的听众作了他的第一个正式的讲座[②]，从周一[③]开始，佩雷尔曼博士将在位于 Stony Brook 的纽约州立大学作另一个系列讲座.

---

① 原题：Celebrated Math Problem Solved，Russian Reports. 作者：Sara Robinson，译自：New York Times，April 15，2003.

② 佩雷尔曼博士于2003年4月7日、9日和11日在麻省理工学院作了3次讲演.——译注

③ 即2003年4月21日.——译注





佩雷尔曼博士拒绝采访,他说宣传为时过早.

Tomasz S. Mrowka 博士,麻省理工学院的一位数学家,参加关于佩雷尔曼博士工作的讨论班已有两个月.佩雷尔曼博士的工作依赖于另一位数学家哈密尔顿所倡导的思想.Mrowka 博士说,到目前为止每次人们提出问题或异议,佩雷尔曼博士总有一个清楚和简洁的回答.

“这是不确定的,但是我们非常严肃地对待它,”Mrowka 博士说道,“显然他思考这个问题很久很深,以至现在很难发现错误.”

庞加莱猜想是法国数学家庞加莱于 1904 年提出的,它是拓扑学中的一个中心问题,而拓扑学是研究当其对象被伸展、扭转或压缩时不改变的那些几何性质的.

地球表面那层薄壳被拓扑学家称之为一个二维球面.它具有这样的性质:其上的每个绳套索可被收紧为一个点.

与此对照,在一个汽车轮胎的表面上,如不割破,一个围绕其中心洞的套索是不能收缩成一点的.

19 世纪以来,数学家们已经知道球面是具有这个性质(即其上的套索可被收缩成一个点——译注)的唯一有界的二维空间,但是高维的情况又如何呢?

庞加莱猜想的一个直观的表述实际上是,三维球面是唯一没有洞的有界三维空间.

“困难在于,当你只能看到它的一小部分时,如何能够知道其整体?”芝加哥大学的一位数学教授 Benson Farb 博士说,“把地球认为平坦的是相当合理的.”

由于出现过许多错误的证明,庞加莱猜想变得很有名.事实上,庞加莱本人也承认,他的猜想的最早形式是错误的.此后,许多数学家曾断言他们得到了证明,但也发现了致命的缺陷.

虽然许多专家说他们对于佩雷尔曼博士的努力感到很兴奋并抱有希望,但他们注意到并非所有的证明都已被写出,另外,即使是最可信赖的研究者也会犯错误,因此他们建议要谨慎.

1993 年,普林斯顿教授安德鲁·怀尔斯博士的情况是,他关于费马大定理的证明被发现有一个严重的漏洞,为了补这个漏洞,怀尔斯博士及其以前的学生 Richard Taylor 博士花费了几个月的努力.

像怀尔斯博士一样,佩雷尔曼博士得到了更好的结果:他实际上证明的是在 1970 年代提出的一个关于三维空间几何的远为广泛的猜想,庞加莱猜想只是其中的一小部分.

佩雷尔曼博士个人的故事可与怀尔斯博士的故事媲美,后者不向其同事泄露而独自在其阁楼上奋战费马大定理.虽然佩雷尔曼博士的早期工作已为其赢得了有才气的数学家的声望,他仍在俄国隐居了最近的 8 年,没有发表任何

文章.

在他去年 11 月张贴的文章中,现年将近 40 岁的佩雷尔曼博士感谢纽约大学的柯朗研究所,位于 Stony Brook 的纽约州立大学和伯克利加州大学,因为他访问这些研究所时的津贴帮助支持了他的俄国的生活.

他宣布,他证明了所谓的几何化猜想(the geometrization conjecture),这是对三维空间几何的一个完全的刻画.

自从 19 世纪以来,数学家们知道,一类二维空间,即二维流形,可以进行分类.

然而,在 20 世纪 50 年代一个俄国数学家证明了:在四维情形,解决这个问题是不可能的,即使在三维情况,问题也是非常非常复杂的.

在 20 世纪 70 年代早期,戴维斯加州大学的一位教授 William P. Thurston 博士猜测:三维流形是由许多齐性片组成,用一种预先描述的方式拼起来的,并证明了他的猜测在许多情形是对的.为此,Thurston 博士赢得了数学界的最高荣誉——一枚菲尔兹奖章.

佩雷尔曼博士的工作——如果正确——则完全描述了三维流形的结构,并且,作为副产品,它解决了庞加莱的著名问题.他的方法利用了所谓的 Ricci 流技巧,这是由目前在哥伦比亚大学的哈密尔顿博士引进的.

Ricci 流是一种平均过程,用来将流形颠簸凹凸处变得均匀平滑.利用 Ricci 流,哈密尔顿博士在某些情形成功地证明了几何化猜想;对一般情形,他勾勒了一个证明方案,然而他碰到的麻烦是,不知道如何控制 Ricci 流的一类奇异现象.

“佩雷尔曼所做的就是,想出了制服这些奇异性的新的和有趣的方法,” Mrowka 博士说道,“他的工作极端地依赖于哈密尔顿的工作,但对于那个计划做出了令人惊奇的新贡献.”

数学家们说,如果佩雷尔曼博士成功地解决了庞加莱猜想,他很可能与哈密尔顿博士分享 Clay 数学研究所的奖金.

即使佩雷尔曼博士的工作没有证明几何化猜想,数学家们说,很清楚,他的工作也将对数学做出本质的贡献.

“无论怎样,都将是令人愉快的,”Mrowka 博士说道,“或者他已经完成了,或者他真正做出了某些重大的进展,我们都将从中学到一些什么.”





数学界对庞加莱猜想的疑似证明众说纷纭①

第八章

数学的一个热门话题：尚未解决的问题不久便会破解．专家们的意见说，即便并非如此，它的自认为的征服者也获得了某些奇妙的东西．

证明庞加莱猜想 = 赢得持久的荣耀．1999 年来，数学家们辛勤却徒劳地陷进了上面这个方程式里．在 2000 年，马萨诸塞州剑桥的 Clay 数学促进会把这个证明的奖金抬高到了 100 万美元．现在，多年来最可信的一位宣称解决了猜想的人正在小心谨慎、一点一点地提出他的证明，在进一步检查之后，它可能就是这最后的一击．也许如此．

上周，在剑桥的麻省理工学院（MIT）的一次被热烈期待的系列讲座中，来自俄国圣彼得堡斯捷克洛夫数学研究所的佩雷尔曼，概述了对此猜想的一个可能证明．想到过去的失败，没有一个人，包括佩雷尔曼本人（它拒绝就此事接受访谈），要宣告已经获胜．但是专家们说，这次努力颇为独特．MIT 的一位拓扑学家 Tomasz Mrowka 说："我们拼命想要弄明白他在这里所做的东西．""很清楚，他已经做出来一些东西．你不会花掉许多时间（去读他的文章）却空手而归．"按圣地亚哥加利福尼亚大学的一位微分几何学家 Bennett Chow 的说法："很清楚，它是一个重大的突破，但是要对他是否已经证明了猜想说任何话都还太早．"

---

① 原载自：Mathematics World Abuzz Over Possible Poincaré Proof．作者：Dana Mackenzie．胥鸣伟，译．李培信，校．译自：Science，2003，1(300)：417．

竞猜游戏:他证明了还是没有证明,开始于 2002 年 11 月 12 日,那时佩雷尔曼在互联网上贴了一份预印本.他神秘地说"我们给出此猜想的一个不拘一格的证明概述".许多读者只不过以为佩雷尔曼勾勒出对此问题一个可能的冲击罢了.但是佩雷尔曼用电子邮件澄清说,他想要公开一个实际的证明.在 2003 年3 月 10 日,他给了第 2 份帖子,已包含了这个工作的更多细节,他还再次说,要完成这项工作.在 MIT 的演讲涵盖了这 3 个分期的材料.

麦比乌斯(Möbius, August Ferdinard, 1790—1868),德国数学家、天文学家

事实上,佩雷尔曼所冲击的问题较法国数学家庞加莱所提出的那个还要宏大.在为现在称为拓扑学的这门学科做奠基工作时,庞加莱怀疑拓扑学是否有足够的工具来区分出最简单的三维对象,即四维蛋的表面.换句话说,一个没有显著特征的三维空间:没有空洞,没有形如麦比乌斯的扭曲,没有手柄,没有边缘,是否它必为一个三维球面?

如果佩雷尔曼的工作被证实了,它也将确认了一个更广的假设.戴维斯加州大学的 William Thurston 的猜想源于 19 世纪德国数学家黎曼的一个里程碑式的定理.黎曼说,任何二维空间(即任意曲面)可以被按摩成在各处均具有相同类型的曲率:或都为负,或都为正,或是平坦.这种"几何化"的曲面具有的曲率越负则它具有的洞越多.一个推论是,一个无空洞的曲面必是正向弯曲的,因此拓扑等价于一个球面.

Thurston 寻求把黎曼的定理带到三维空间,即他的"几何化猜想".因为三维空间远较二维复杂,数学家们不能按黎曼那样把它们熨烫成常曲率.在 1970 年代后期,Thurston(那时他在 Princeton 大学)转而提出他们可以做成几乎一样好的事.他猜想,在适当的地方进行切割,他们可以把任何三维空间分成若干片,它们可以转换成 8 种非常标准的几何之一,这些几何涉及的范围从双曲的(负向弯曲)到球面的(正向弯曲).上述猜想如果为真,则几何化猜想就会给数学家们某种"周期表",用来对三维空间分类.它也会立即解决了庞加莱猜想.因为 7 种非球面的几何中每一个都将留下泄露其真面目的拓扑指纹,于是那个没有识别记号的空间就会是球面.证毕.但是如何证明此猜想呢?

1980 年代早期,哈密尔顿(现在 Columbia 大学)提出,数学家可以诱导一个三维空间实现自身几何化,方法是促使它"流"向一种标准的几何,这很像是穿过铁栅的热流.1988 年哈密尔顿利用他的"Ricci 流"重新证明了二维曲面时的黎曼定理.

但是在三维情形,哈密尔顿遭遇到了一个问题.一个空间的不同部分的流动会扩张它们的不同几何,使得它们之间的边界拉伸成无限薄的"颈部".就像哈密尔顿的同事,哈佛大学的丘成桐所指出的那样,这些颈部标出了那些需要数学家进行"外科手术"的地方,这种手术是 Thurston 猜想所需要的.但是如果颈部过快折断或是形成了棘手的形状,包括一种特别麻烦的称做"雪茄"的形





状,手术就会失败.进一步,又是什么能避免不会像好莱坞明星们的整容手术那样,无休止地进行下去?

丘成桐和哈密尔顿两人都和这些问题搏斗了数年.佩雷尔曼使用了许多由他们发展起来的技巧,但是却加进了一个关键的思想.他分辨出一类"熵",它随空间流动而增加.用对 Ricci 流给出某种意义下方向的方法,佩雷尔曼的熵有助于保持一个空间朝向几何化运动.更微妙地,它也让佩雷尔曼控制住了塌陷区域的大小和形状.

虽说不是全部但的确有许多专家似乎相信佩雷尔曼已经掐灭了雪茄并制服了狭窄的颈部,然而他们对他能否控制住手术次数则信心不足.注意到其他对庞加莱猜想所尝试的证明都在相似的缺失步骤上失足,丘成桐警告说,那可能证实为一个致命的缺陷.即使佩雷尔曼没有达到他的主要目标(和那个 100 万奖金),Bennett Chow 说,他的工作代表了在理解 Ricci 流方面的巨大进步."除去在现实世界我们知道山有多高外,它像是在登山,哈密尔顿所做的是登上了一个难以置信的高度,远超过人们所想到的高度,而佩雷尔曼则从哈密尔顿离开时的高度开始并已经登到了更高的地方但我们依然不知道这座山到底有多高."

中　编

三维空间与拓扑学



空间为什么有三维?[①]

第九章

1 “拓扑学”和连续统

几何学家通常在两类几何学之间做出区分,他们把第一类称为度量几何学,把第二类称为射影几何学.度量几何学以距离概念为基础;在度量几何学中,当两个图形“全等”(在数学家赋予这个词的意义上)时,则它们被认为是等价的.射影几何学以直线概念为基础.因为在射影几何学中,认为两个图形等价并不一定要它们相等,只要它们通过射影变换彼此对应(即一个是另一个的射影)就足够了.第二类几何学往往被称为定性几何学:若与第一类几何学相比较,它的确是这样.显然,在射影几何学中,度量和量并不起什么重要的作用.然而,也不完全如此.直线不是纯粹定性的;在没有做出某种度量或者在没有使所谓的直尺(一种度量工具)沿一条线移动的情况下,就不能断言这条线是直线.

但是,还有第三类几何学,在这类几何学中,量被完全排除了.它纯粹是定性的,这就是拓扑学.在这个学科中,可以通过连续变形使一个图形与另一个图形对应,从而两个图形在任何时候都是等价的,不管支配这种变形的规律是什么,只要保持连续性就行.于是,圆等价于椭圆,甚至等价于任何类型的闭曲线,

① 原载自:庞加勒.最后的沉思[M].李醒民,译.范岱年,校.北京:商务印书馆,1995:30-51.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

但它与线段不等价，因为线段不是闭合图形．球面等价于任何曲面，但是它不等价于圆环面，因为在圆环面上有一个洞，而球面上却没有．让我们设想任何一类图样，一个笨拙的制图员描画这个图样的复制品．比例被歪曲了，用颤抖的手画出的直线歪歪扭扭，结果成了不成比例的曲线．从度量几何学的观点来看，甚至从射影几何学的观点来看，这两个图形都不是等价的；但是，与之相反，从拓扑学的观点来看，它们是等价的．

对于几何学家来说，拓扑学是很重要的科学．拓扑学导致了一系列定理，这些定理像欧几里得的定理一样密切相关；正是从这组命题出发，黎曼构造了一种最著名的、最抽象的纯粹分析理论．为了说明它们的本性，我将引用其中的两个定理：(1)平面上的两个闭曲线相交于偶数个点；(2)如果一个多面体是凸多面体（这就是说，如果不把它一切为二就不可能在它表面上描绘一个闭合线），那么它的棱数等于顶点数加面数减去2；当多面体的面和棱是曲面和曲线时，这依然是正确的．

这就是拓扑学使我们如此感兴趣的东西．正是在这门学科中，几何学直觉确实起着作用．在度量几何学的定理中，当运用能力是由这种直觉组成时，那正是因为在无视一个图形的定性性质时，也就是说，在忽视研究那些严格地属于拓扑学的性质时，便不可能研究它的度量性质．人们常说，几何学是一门关于粗制滥造的图形的正确推理的艺术．这不是冷嘲热讽，而是值得思考的真理．但是，什么是粗制滥造的图形呢？刚才提到的那位笨拙的制图员所能画出的图形就是这类图形．他或多或少公然地歪曲了比例；他把直线乱画为锯齿形；他的圆好像土堆一样难看．但是，所有这一切无关紧要；它无论如何不会使几何学家烦恼；这并不妨碍他正确地推理．

但是，缺乏经验的画图者必然不用开曲线描绘闭曲线，或者不用没有公共点的三条直线描绘相交于一点的三条直线，或者不用完整的曲面描绘有洞的曲面．在那种情况下，这位画图者的图画毫无用处，推理也变得不可能了．直觉不会受到图画中仅对度量几何学和射影几何学有意义的缺陷的妨碍．然而，只要这些缺陷涉及拓扑学，直觉将变得不可能．

这种十分简单的观察指出几何学直觉的真实作用；几何学家需要画图形，至少需要形成它们的思想图像，从而便利了这种直觉．现在，如果他尽量减小这些图形的度量性质和射影性质的重要性，如果他仅仅专注于它们的纯粹定性的性质，那么唯有几何学直觉在这里真正起作用．我并不是说度量几何学是建立在纯粹逻辑的基础上，或者其中没有直觉真理的地位．但是，它们是另一类直觉观念，类似于在算术和代数中起主要作用的直觉观念．

拓扑学的基本命题是：空间是三维连续统．我已经在其他著作中讨论了这





个命题的起源,但却是以极为简略的方式讨论的,为了阐明某些观点,再次更详细地考察一下它,在我看来并非是毫无意义的.

空间是相对的;所谓相对空间,我不仅意指在我们没有注意到的情况下,我们可以转移到空间的另一个区域(这是我们真正遇到的事情,因为我们并不觉察到地球的平动);我不仅意指,一切物体的所有维数在我们不能知道其变化的情况下能够成比例地增加,倘若我们的测量仪器经受到同样的变化的话;而且我也意指,空间能够按照某个任意的规律变形,假使我们的测量仪器也按照这个同样的规律变形的话.

这可以是任何变形,但变形必须是连续的;也就是说,它必须是使一个图形变换为从拓扑学观点来看是等价的另一个图形的那些变形之一.当空间被认为是独立于我们的测量仪器时,空间从而既不具有度量的性质,也不具有射影的性质;它只有拓扑的性质(也就是说,仅具有在拓扑学中所研究的性质).它是无定形的,也就是说,它并非不同于人们通过无论什么连续性的形变能够从它得出的任何空间.我将用数学语言加以解释.在这里有两个空间 E 和 E';E 中的点 M 对应于 E' 中的 M';点 M 有直角坐标 x,y,z;点 M' 具有 x,y,z 的三个任何连续函数作为直角坐标.从我们所谈到的观点看来,这两个空间并没有什么不同.

我们测量仪器的功能,尤其是固体的作用如何给人的智力提供更完满的决定和组织这种无定形空间的机会,它怎样容许射影几何学画直线网络,怎样容许度量几何学测量这些点之间的距离群的基本概念在这个过程中起什么根本性的作用,我在其他著作已经对此作了详细的解释.我认为所有这些论点都已得到确认,我不需要再重复这些了.

在这里,我们只关心在拓扑学中所考虑的无定形的空间,即独立于我们测量仪器的唯一的空间;它的基本性质——我是要说它的唯一的性质——是三维连续统的性质.

2 连续统和截量

可是,什么是 n 维连续统呢?它与维数较大或较小的连续统怎样区别呢?让我们首先回顾一下康托的学生最近得到的一些结果吧.在直线上的点和平面上的点之间,或者更一般地说,在 n 维连续统上的点和 p 维连续统上的点之间有可能建立一一对应关系.倘若我们不受平面上两个无限邻近的点对应于直线上两个无限邻近的点这个条件(即连续性条件)的约束,那么这就是可能的.

因此,有可能用这样的方式使平面发生变形而得到直线,只要这种变形不

是连续的.另一方面,用连续的变形则不可能这样.于是,维数的问题与连续性概念密切相关,而对于任何想要排除这一概念的人来说,那是没有什么意义的.

为了定义 n 维连续统,我们首先有解析定义:n 维连续统是 n 个坐标的集合,也就是说,是能够各自独立变化的,而且假定所有的实值满足某些不等式的 n 个量的一个集合.这个定义从数学的观点来看尽管没有缺点,但是无论如何不能使我们完全满意.在连续统中,各种坐标可以说并非相互毗连;它们在它们自身之中联系起来,以至形成一个整体的各个方面.在空间研究的每时每刻,我们实现的就是所谓的坐标变换.例如,我们实现直角坐标系变换,要不然我们变换到曲线坐标.在研究另一个连续统时,我们也实现坐标变换;也就是说,我们用 n 个坐标的无论什么样的 n 个连续函数代替 n 个坐标.对于我们之中不是从刚才提到的解析定义出发,而是从某个更深奥的来源出发而导出 n 维连续统概念的人来说,这一操作是很自然的;我们感到,那些在连续统中是本质的东西并没有变化.另一方面,对于那些仅仅从解析定义了解连续统的人来说,这一操作无疑是合理的,但却是奇异的,未经证明的.

最后,这个定义尽量减小了连续统概念的直觉起源和这一概念所包含的一切丰富思想的重要性.它像那些从数学"算术化"以来在这门科学中变得如此频繁的定义那样反复出现.从数学的观点来看,我们所说的这些定义是没有缺点的,但是它们却不能使哲学家满意.它们用由比较简单的材料组成的结构代替被定义的对象和这个对象的直觉概念.因此,很容易看到,用这些材料可以有效地形成这个结构,但我们同时看到,要做出更多的东西同样是可能的.未被揭示出来的是:为什么用这种方式而不用另外的方式来组合这些材料,其中有什么深刻的原因?我的意思并不是说,数学的这种"算术化"是不受欢迎的;我说它并非包罗万象.

我将把维数的确定建立在截量概念的基础上.首先,让我们考虑一条闭曲线,即一维连续统.如果我们在这条曲线上取任意两个我们将不容许我们自己通过的点,那么该曲线将被截为两部分,不可能从一部分到另一部分,因为我们虽然还在这条曲线上,但是却不能通过被排除的点.另一方面,让我们考虑一个闭曲面,它形成一个二维连续统.在这个曲面上,可以取一个、两个或任意数目的被排除的点.该曲面并不因为这样就被分为两部分;在这个曲面上,可以从一点到另一点,而不会遇见任何障碍,因为总可以绕过被排除的点.

可是,如果我们在曲面上画出一条或多条闭曲线,如果我们把它们看做是不可逾越的截量,那么该曲面就能够被分为几个部分.

现在,让我们考虑空间的情形.我们既不能禁止通过某些点,也不能禁止通过某些线来把空间分为几个部分,这些障碍总可以绕过去.必须禁止越过某些





面,即某些二维截量.这就是我们说空间具有三维的原因.

我们现在知道,n 维连续统是什么.当一个连续统能够借助于一个或多个本身是 $n-1$ 维的截量被分为许多区域,则该连续统具有 n 维.这样,n 维连续统用 $n-1$ 维连续统来定义.这就是递归定义.

在这个定义中,什么东西给我以信心呢?什么东西向我表明观念实际上如何自然而然地在人们的头脑中产生呢?它首先就是,许多基本读物的作者并无意于恶作剧,但在他们著作的开头部分却做出了类似的事情.他们把体积定义为空间的部分,把面定义为体积的边界,把线定义为面的边界,把点定义为线的边界;此后他们停顿下来,其类似性是明显的.遵循这种定义,我们在拓扑学的其他部分重新发现截量的重要作用.例如,根据黎曼的观点,是什么东西把圆环面与球面区别开来呢?正是这样的事实:我们不能在球面上画一条闭曲线而又不把球面分为两部分,可是却存在着不把圆环面分为两部分的闭曲线,为了保证人们分开圆环面,必须做出没有公共点的两个闭截量(闭曲线).

还留下另一个值得考察之点.我们刚才考察的连续统是数学连续统;它们的每一个点都是独特的东西,绝对不同于其他点,而且绝对不可分.由我们的感觉所直接揭示的连续统,我称之为物理连续统,它们都是有差别的.支配这些连续统的规律是费希纳(Fechner)定律,我将剥去通常套在它身上的华丽的数学外衣,以便把它还原到作为它的基础的实验数据的简单项.根据估计,有可能分辨出一个 10 g 的砝码和一个 12 g 的砝码的差别,但恐怕不可能分辨出一个 11 g 的砝码一个 10 g 的砝码或 12 g 的砝码的差别.更一般地,可以有这样两个感觉集合:我们在没有分辨出一个集合或另一个集合与第三个集合的差别的情况下就可以分辨出它们二者的差别.根据这一假定,我们能够设想这样一个感觉集合的连续链,它们中的每一个都无法与相接的一个区别开来,尽管链的两端却能够很容易地加以分辨.这将是一维的物理连续统.我们也可以设想较复杂的物理连续统.这些物理连续统的元素将又是感觉的集合(但是我更喜欢用比较简单的词——元素).另外,什么时候我才能说,相似元素的系统 S 是物理连续统呢?无论任何时候,我都能够把它的任意两个元素看做是一个连续链的两个末端,该链类似于我刚刚叙述过的链,它的所有元素都属于 S.因此,如果可以用不离开曲面的一条连续的线联结该曲面的任何两个点,那么该曲面就是连续的.

我们能够把截量的概念推广到物理连续统,从而决定它们的维数吗?我们显然能够这样做.让我们排除 S 中的某些元素以及所有不能与它们区分的元素.这些受到限制的元素完全可以是有限的数目,要不然就能够通过它们的结合形成一个或多个连续统.这些有限的元素的集合将组成一个截量;在形成这

一截量后，所发生的情况是，我们可以把连续统 S 分为几个别的连续统，这时再也不能通过连续链从 S 中的任何元素到任何其他元素中去，这个链的元素无法与该截量的任何其他元素相区别.

因此，通过把我们自己限制到有限数目的元素之内，从而能够被截的物理连续统将具有一维；如果一个物理连续统能够借助于本身是 $n-1$ 维的物理连续统的截量来分割，那么它将具有 n 维.

3 空间和感觉

问题似乎被解决了；我们也许只需要把这个法则应用于作为空间的粗糙图像的物理连续统，或者应用于对应的数学连续统——它是物理连续统的精制的图像，是几何学家的空间.但是，那是一种假象；如果我们由以推知空间的物理连续统是直接通过感觉揭示给我们的，那么一切也许是幸运的：然而，事实却远非如此.

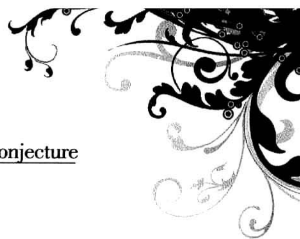
让我们看看，从我们的大量感觉中实际上是怎么有可能推导出物理连续统的呢？物理连续统的每一个元素都是感觉集合；首先考虑一下同时的感觉的集合，即意识的状态，这是最简单的集合.然而，我们的每一个意识状态是一种极其复杂的东西，以至于我们从来也不能指望看到两个意识状态变得不可区分.可是，为了构造物理连续统，从以前已说过的情况来看，基本的问题是，它们的两个元素在某些情况下能够被看做是不可区分的.可是，我们永远也不能说：我不能把我目前的思想状态与我前天同一时刻的思想状态区分开来.

因此，我们有必要通过积极的思想操作，通过忽略两个意识状态的差别，从而一致认为二者是等价的.例如，我们可以忽略某些感官的感觉，这将是最为简单的.我已经说过，我无法分辨一个 10 g 的砝码和一个 11 g 的砝码的差别.可是，情况也许是，如果我不断地实验，那么一个 10 g 的砝码所引起的压力感觉被各种不同的嗅觉和听觉伴随着，当用一个 11 g 的砝码代替一个 10 g 的砝码时，这些各种各样不同的感觉变化了.正因为我忽略了这些特异的感觉，我才能够说，两个意识状态是不可区分的.

有可能规定更复杂的条件，也有可能以不仅把同时的感觉的集合，而且把相继的感觉的集合即感觉系列看做是我们的连续统的元素.接着，有必要规定基本的条件，而且为了认为连续统两个元素是等价的，有必要指明二者必须具有的共同特性(不管它们是同时的感觉的集合还是相继的感觉的集合).

于是，在定义物理连续统的场合，有必要做出双重选择：第一，选择作为这个连续统的元素的同时的或相继的感觉集合；第二，选择定义两个元素必须被





认为是等同的情况的基本条件.

为了得到空间,必须怎样进行这种双重选择呢?我们能够满足于考虑同时的感觉的集合或者有必要考虑感觉系列吗?特别是,我们能够以由于忽略某些感官的知觉而形成的最简单的和最自然的基本条件为满足吗?否!

这样的否定是不可能的;我们不能从我们的感觉中选择出那些将向我们传达空间概念并且只传达空间概念的感觉.没有一种感觉不借助于其他感觉就能够向我们传达空间概念;也没有一种感觉不传达大量与空间毫无关系的东西.

例如,我们分析一下所谓接触的知觉,这是我们觉察到的知觉.经验告诉我们,如果我们用两个大头针接触我们的皮肤,倘使它们相距足够远,那么我们的意识就能够分辨出这两个大头针,如果使它们相互靠得很近,我们就无法在二者之间做出区分了.而且,区分它们的最小距离依据身体部位而变化.我们通常说,皮肤被分为各个部位,每一个部位都是同一感觉神经的管辖范围;如果两个大头针扎入同一部位,那么只有单根神经受到刺激,我们只意识到一个大头针;但是,换一种情况,如果它们扎入两个部位,结果影响到两根神经,我们便觉察到两个大头针,这并不完全令人满意;我们无法用这种方式发现物理连续统的特性.让我们设想一下,我们改变两个大头针的位置,而使它们已经很小的距离保持恒定.由于这个距离很小,可以发生下述情况:两个大头针将扎入同一部位,结果只产生一个知觉.但是,如果我们一点一点地改变它们的位置,而不改变它们的距离,在某一瞬间,将出现这样的情况:它们中的一个将扎入该部位之外,而另一个还处于该部位之内.在此瞬间,我们应当感觉到两个大头针,但我们所观察到的情况并非如此.我们不可能用这种方式推断出物理连续统的概念,但是却可以推断出由像有那么多部位那么多的独特情况所形成的离散集的概念.最好是姑且承认,大头针的接触不仅影响最近的神经,而且也影响相邻的神经,而当距离增大时,其强度亦随之减小.因此,让我们设想,我们正在把两个大头针接触的作用进行比较.如果两个大头针的距离很小,那么同一神经受到作用;某一个大头针对于同一神经的刺激强度将无疑是不同的,但是这种差别太小了,以至于按照费希纳的一般法则也难以分辨出来.如果一根神经受到大头针 A 的刺激而没有受到 B 的刺激,那么它仅仅是受到大头针 A 的轻微刺激,这个刺激将低于"意识阈限".因此,两个大头针的影响将是不可区分的.

这样,我们有了我们为构造物理连续统所需要的一切;我们只要使两个大头针沿着我们皮肤的表面移动,我们只要注意在哪一种情况下我们的意识能分清它们.我们已略去了(那是我上面所提到的作为我们基本的条件的东西)大量的事实:每一个感觉网络的刺激强度、大头针在皮肤上所施加的或大或小的压力、接触的性质.触觉揭示出了所有这些事实,但是我们排除了它们,以便只保

持其特性是几何学的那些事实.这样一来,我们推断出空间概念了吗？没有;首先,这样构造出的连续统像皮肤本身的表面一样只有二维.其次,我们十分清楚地知道,我们的皮肤是可动的,皮肤上的特定点并不总是对应于空间的特定点;当我们的身体变形时,皮肤上两点之间的距离就要发生变化.毫无疑问,软体动物正是用这种方式想象空间的,但是这与我们的空间概念无关.

同样的情形对视觉也是真的;照射到视网膜两点上的两束光,根据这两点的距离是大还是小,要么给我们以两个光斑的印象,要么只给我们一个光斑的印象.这相当于上述的两个大头针;我们能够忽略光的颜色和强度,利用它们构造物理连续统;这个物理连续统正像视网膜的表面一样,将具有二维.第三维是通过眼睛的双目视觉的会聚作用引入的,这就是所谓的视觉空间(Visual Space).它高于触觉空间(Tactile Space),首先是因为我们怀着一点善意给它以三维,其次是因为视网膜无疑是可动的,而从固体的意义上讲,皮肤却在所有方向上都是柔韧的.于是,我们被诱使说,真实的空间存在于我们企图确定我们所有的感觉起源的地方.这还不能使人满意.不仅眼睛是可动的,以至于空间的特定点并不总是对应于视网膜的特定点和眼睛的特定会聚度;而且这也无法解释,为什么第三维如此明显地与已经引入的其他二维不一致,也无法解释为什么盲人的几何学和我们的相同.

如果我们希望把视觉空间和触觉空间结合起来,那么将有五维而不是三维或二维;将依然存在着用什么过程解释五维能够简化为三维的任务;如果我们希望把其他感觉引入这种结合之中,那么维数将进一步增加.

还要用几句话来解释,为什么触觉空间和视觉空间是同一个空间.

4 空间和运动

因此,情况似乎是,我们不能通过考察同时的感觉的集合来构造空间,我们必须考虑感觉系列,总是有必要再次提到我前面已经说过的东西.某些变化表现为位置的变化,另一些变化表现为没有几何学性质的状态的变化,这究竟是为什么？为此,我们必须首先区分外部变化和内部变化;外部变化是非随意的,它们并不被肌肉感觉所伴随;内部变化是我们身体的运动,我们可以把它们与其他变化区别开来,因为它们是随意的,并被肌肉的感觉所伴随.内部变化能够矫正外部变化,例如我们以这样的方式用我们的眼睛跟踪运动着的物体,使它的映象总是返回到视网膜的同一点上.可以被这种矫正感受的外部变化是位置变化;如果它不能被这种矫正感受,它就是状态变化.

从定性的观点看来是完全不同的两种外部变化,如果能够用相同的内部变





化来矫正它们,那么它们就被认为是对应于同一位置变化.也可以这样说,如果两个内部变化能够矫正相同的外部变化,那么它们就能由毫无共同之处,但是却对应于同一位置变化的肌肉感觉系列组成.这就是当我们说,有许多路线能够从一点引到另一点时,我们用通常的用语所表达的意思.

因此,重要的是,为了到达特定的物体,必须做的就是动作.对于我们来说,这些动作的意识无非是伴随它们的肌肉的感觉集合.

由此推断,某一物体与我的一个手指接触;比方说,与我右手的食指接触.从这一事实我经验到触觉 T;同时,我从这个物体经验到视觉 V.当把该物体移开时,感觉 T 逐渐消失,视觉 V 被新的视觉 V' 代替;这是一种外部变化.假定我希望通过复原感觉 T,即使我的食指再次接触该物体,来部分地矫正这一外部变化.为了做到这一点,我必须完成某些动作,对我来说,这些动作通过肌肉感觉系列 S 表示出来.我知道,这是因为我或我的祖先的大量经验告诉我,当感觉 T 消失而视觉从 V 变到 V' 时,可以通过对应于该系列 S 的运动来复原感觉 T.我同样清楚地知道,对我来说,我通过不用系列 S,而用另外的系列 S' 或 S'' 描述它们自身的其他功作而能够得到相同的结果.

所有这些肌肉感觉系列 $S,S',S'',\cdots$ 或许没有共同的元素;我之所以比较它们,是因为我知道,它们中的任何一个在视觉 V 变为 V' 的每一时刻都能够复原感觉 T.用我们通常的语言,已经通晓几何学的我们将说,对应于肌肉感觉系列 S,S'',S'' 的各种动作系列有这样的共同之处:在它们任何一个中,我们食指的初始位置和最终位置依然相同.其他每一情况可能不同.

这样,我未被引导去区分这些不同的系列 $S,S',S'',\cdots$,也没有把它们视为单一的感觉.我不想去区分与这些系列差别过小的肌肉感觉系列.届时,我将有构造物理连续统的方法.事实上,我已选出这个连续统的元素,它们是肌肉感觉系列,而且我有了"基本的条件",这些条件告诉我,在哪一种情况下,这些元素中的两个必须被视为是等同的,正是这种连续统有三维.

可是,这并非一切.我们刚刚定义了一个是真实空间的连续统;正是这个空间,被看做是用我的一个手指描述的.但是,我有几个手指(而且从与我有关的观点来看,所有我的皮肤上的点都可以视为手指).我的不同的手指将描述相同的空间吗?是的,毫无疑问,可是这意味着什么呢?这意指的是性质的集合,用通常语言不容易描述它,如果容许我用某些符号,我可以尝试解释它.我将考虑两个手指,并称之为 α 和 β;手指 α 比如说是右手的食指,我们为定义系列 S,$S',S'',\cdots$ 曾使用过它.然后我将写出

$$S \equiv S' \pmod{\alpha}$$

这意味着,如果对应于 S 的动作恢复用手指 α 所经验到的触觉,那么同样的情

况对于对应于 S' 的动作也是真的，反之亦然.类似地，我将写出

$$S_1 \equiv S'_1 (\bmod \beta)$$

来描述下述事实：如果对应于 S_1 的动作恢复用手指 β 所经验到的触觉，那么同样的情况对于对应于 S'_1 的动作也是真的.

在作这种推断之后，我将假定存在着两个特定的肌肉感觉系列 s 和 s_1，它们是以下述方式被定义的：我将设想，手指 β 由于与一个物体接触而经验到触觉.通过完成对应于 s 的动作，这一感觉将消失.可是，最终将是手指 α 经验到触觉.我通过经验知道，在这些动作之前，在手指 β 感觉到接触的每一时刻（或者，至少几乎在每一时刻），都会发生这种情况.（我之所以说几乎，是因为要相继发生，便要求该物体在这一时间间隔内不运动）用我们通常的语言（这种语言对我们来说比较清楚，但是我不敢使用它，因为我讲的是还不具有任何几何学知识的人），我可以说，对应于 s 的动作把手指 α 引到手指 β 原先占据的位置.对于 s_1 来说，相反的情形将是真的；对应的动作将把手指 β 引向手指 α 原先占据的位置.如果这两个系列 s 和 s_1 存在关系

$$S \equiv S' (\bmod \alpha)$$

将导致作为结果的下述关系

$$s + S + s_1 \equiv s + S' + s_1 (\bmod \beta)$$

如果我们回想一下符号的意义，我们便会立即相信上述关系，我们还可以从它毫无困难地推出，由 α 和 β 产生的两个空间是同构的，特别是，它们有相同的维数.

如果系列 s 和 s_1 不存在，那么同样的情况便不可能为真.事实上，让我们设想，不可能找到一个动作系列，这个系列将在手指 β 与物体接触的感觉上引起手指 α 与同一物体接触的感觉——肯定地或者至少是几乎肯定地——这时我们应当如何推理呢？我们可以说，手指 β 感觉到物体没有位于空间同一点，它感觉到物体隔着一段距离；另一方面，每次手指 β 之所以感觉到该物体，那可能是因为物体处于空间中的同一点 A.因而必须存在着把手指 α 引向点 A 的动作系列.由于物体处于点 A，手指 α 应该能够感觉到物体，这件事总是应该发生.因此，如果我们假定不存在具有这一性质的动作系列，那么我们就必须承认，手指 β 感觉到在一段距离之外的接触；换句话说，为了确定物体在空间的位置，对于该物体来说，被手指感觉到并不充分；最后，这也就是说，空间必定比用手指按照我们描述过的方式产生的物理连续统有更多的维数.

例如，我将假定，空间具有四维，我将用 x, y, z, t 来表示四个坐标.我将假定，手指 β 每时都感觉到与物体接触，此时三个坐标 x, y, z 对于手指和物体都是相同的，而不管第四个坐标可能是什么；而且，手指 α 每时都感到与物体接





触,此时三个坐标 x,y,t 对于物体和这个手指都是相同的,而不管坐标 z 可能是什么.在这些条件下,让我们把我们的法则用来构造由 β 产生的物理连续统;我们将发现,它只有三维,这三维对应于三个坐标 x,y,z,坐标 t 不起任何作用.按同样的方法,由 α 产生的物理连续统有三维,它们对应于 x,y,t.但是,我们不能够找到对应于这样的肌肉感觉系列 s 的动作系列,以至于对 α 的接触感觉肯定地随着对 β 的接触感觉.

事实上,设 x_1,y_1,z_1,t_1 是物体的坐标;手指 β 在动作之前的坐标是 x_0,y_0,z_0,t_0;手指 α 在动作之后的坐标是 x_0',y_0',z_0',t_0'.我们将用下述写法表示手指 β 在动作之前感觉到接触这一事实

$$x_0=x_1,y_0=y_1,z_0=z_1 \quad ①$$

我们将用写法

$$x_0'=x_1,y_0'=y_1,z_0'=z_1 \quad ②$$

表示 α 在动作之后感觉到接触的事实.

因为 s 存在,我们必然能够以这样的方法来选择 x_0,y_0,z_0,t_0 和 x_0',y_0',z_0',t_0',使得关系式①能够导致关系式②,而不管 x_1,y_1,z_1,t_1 可能是什么.很清楚,这是不可能的.恰恰是不可能形成 s 的这一点在这种情况下向我们揭示出,空间应当有四维,而不像 β 产生的物理连续统那样只有三维.

再者,如果我们引入视觉,那么我们实际上会观察到某种类似的事情.让我们考虑视网膜上的一点;我们能够赋予它像我们的手指 α 和 β 一样的作用.我们能够设想必然使物体的映象反映到视网膜的点 γ 上的动作系列或肌肉感觉 S 的对应系列.我们能够利用这个系列,以便定义类似于由 α 或 β 所产生的物理连续统.这个连续统将只有二维.但是,我们不能构造类似于 s 的系列,也就是说,不能构造这样一个动作系列:作为在点 γ 感觉到的视觉结果,该动作系列肯定引起手指 α 感觉到的触觉.换句话说,因为我们观察到物体的映象在 γ 发生,就是说我们能够确定该动作必然引导我们的手指与这个物体相接触,这没有充足的理由.我们缺乏一项关于物体的距离的资料.这就是为什么我们说,视力在一段距离之外起作用,空间有三维——比 γ 产生的连续统多一维.

从这个简短的叙述中,我们看到,导致我们把三维赋予空间的实验事实是什么.考虑到这些事实,在我们看来,赋予空间以三维,而不是四维或二维,更为方便一些.但是,"方便"这个词不可能有足够强的说服力.把二维或四维赋予空间的人会发现他自己在像我们这样一个世界的生活斗争中是很不利的.这实际意味着什么呢?让我再次提到我的符号,例如全等

$$S \equiv S' \pmod{\alpha}$$

它的意义我在上面已经解释过了.把二维赋予空间就得要承认我们自己并不承

认的类似的全等. 这时,我们便被导致用做不到的动作 S' 来代替能顺利进行的动作. 相反地,把四维赋予空间,就会排斥我们自己承认的全等. 因此,我们就会剥夺我们自己用其他动作 S' 代替动作 S 的可能性,尽管 S' 这些动作同样有效,并且在某些情况下,它也许还会带来特殊的好处.

5 空间和自然界

可是,问题能够从完全不同的观点提出来. 直到现在,我们采取的观点纯粹是主观的,纯粹是心理学的,或者如果我们希望的话,也可以说是生理学的. 我们只考虑了空间与我们的感觉的关系. 另一方面,我们能够采取物理学的观点,我们可以问我们自己,是否能把自然现象定域在其他空间内,而不是定域在我们自己的空间内,例如定域在二维或四维空间内. 物理学向我们揭示的规律是用微分方程描述的,在这些方程中包含着某些质点的三个方程,用其他方程,例如包含具有四个坐标的一些质点的方程,描述同一规律是不可能的吗? 或者,这也许是可能的,但是由此得到的方程却较不简单? 最后,或者它们却是如此简单,而我们却要完全抛弃它们,只是因为它们扰乱了我们的思想习惯?

当我们说用其他方程描述同一规律时,我们意味着什么呢? 让我们考虑两个世界 M 和 M'. 我们能够在这两个世界中发生的或可能发生的现象之间建立这样一种对应关系,使得对于第一个世界的每一个现象 φ 对应于另一个世界完全确定的现象 φ' 也可以说是 φ 的映象. 从而,如果我假定,在遵循支配世界 M 的规律的情况下,现象 φ 的必然结果是某个现象 φ_1,作为 φ 的映象的现象 φ' 的必然结果,在遵循支配世界 M' 的规律的情况下恰恰是现象 φ_1 的映象 φ_1',那么我们就能够说,这两个世界服从同一规律. 现象 φ 和 φ' 的质的本性对我们来说并不怎么重要;“平行关系”是可能的这一点就有充分的理由了.

而且,事实上,现象的质的本性只是我们的感官关心的东西,我们已经同意采取超心理学的观点,因此可以忽略我们感官的感觉,而只把注意力放在现象的相互关系上. 事实上,例如当物理学家用仅看到运动质点的分子运动论的气体来代替我们通过经验所熟知的产生压力和热感觉的气体时,或者用以太振动来代替我们经验到的光和光产生的色感时,他就是这样做的.

只要考虑一个简单的例子,即天文学现象和牛顿定律的例子就足够了. 我们观察到的东西不是天体的坐标,而仅仅是它们的距离. 因此它们的运动规律的通常表达式是这些距离和时间的微分方程. 现在,空间两点之间的距离是一个已知的这两点的坐标的单叶函数. 让我们通过在微分方程中用这种函数代替每个距离,来变换我们的微分方程. 这时我们便有它们的通常形式的方程,天体





的坐标本身包含在这种形式中.

但是,我们可以用其他函数来代替这些距离.从而能够得到这些方程的其他形式.从与我们有关的观点来看,所有这些形式是同等合理的,因为它们服从现象中的“平行关系”.让我们设想天体以这样的方式处于四维空间中,它们每一个的位置不再由三个坐标,而是由四个坐标来确定.接着,让我们在方程中用两个天体的八个坐标的无论什么函数来代替迄今我们视为描述这两个天体之间距离的量.在通常的四维空间中,根本没有必要使这个函数是描述两点之间的距离的函数;它可以是无论什么函数,因为这不会违反“平行关系”.

从而,我们将得到我们方程的一种形式,在这种形式中,涉及天体在四维空间的坐标.这将是以四维空间假说为基础的天文学定律的新表述,这一表述不会与该定律背道而驰,因为它服从“平行关系”条件.不管怎样,这样得到的方程不用说远没有我们通常的方程简单,这一点是很清楚的.

毋庸置疑,同样的情况对于物理学规律来说也是真的.存在着一般的理由,使得它应当如此吗?即在所有的物理学分支中,是有关三维性的假说给这些方程以其最简单的形式吗?这个理由与我在这篇文章的第一部分所提到的东西,与绝对地迫使一切人相信三维性的东西,或者在人们处于生活斗争不利地位的困境下迫使人们好像相信三维性似的那样行动的东西有任何关系吗?

在这里,有必要简短地说一点题外话.例如,让我们再次把我们通常的空间归于我们的创造者.我们说空间是相对的,这意味着物理学定律在这个空间的所有部分是相同的;或者,用数学语言来说,就是描述这些规律的微分方程不依赖于坐标轴的选择.

如果我们考虑一个完全孤立的系统,那么这没有什么意义;不可能观察这个系统的点的坐标,而只能观察它们的各自距离.观察将不会告诉我们,这个系统的性质是否取决于该系统在空间的绝对位置,因为这个位置是不可观察的.

如果系统不是孤立的,事情也不可能是这样(如果我们希望以严格的精确性进行论证的话),因为在没有考虑到外部物体作用的情况下,不可能描述支配这个系统的规律.可是,却存在着几乎孤立的被其他物体包围的系统,这些物体要近到足以被看得见,然而又远到难以感觉到它们的作用力.对于与恒星有关的我们的地上世界来说,所发生的情况就是这样.因此,我们可以阐明这个地上世界的规律,就好像恒星不存在一样,但我们仍可以把这个世界与完全确定的并与这些恒星不变地联系在一起的坐标系关联起来.所以,经验告诉我们,坐标系的选择无关紧要,当进行坐标变换时,方程不会不成立.正如我们知道的,坐标轴的可能变换的集合形成一个六维群.

让我们撇开我们通常的空间不谈,让我们用在服从现象“平行关系”的意义

上是等价的其他方程来代替我们的方程.每当我们涉及近似孤立的系统时,将存在极其普遍的事实和将保持不变的不变性特性;将存在不会使方程不成立的变换群.这些变换将不再具有坐标轴变换的含义,它们的含义能够是无论什么东西,可是这些变换所形成的群必须始终与我们刚刚提到的六维群保持同构.没有这一点,就不会有任何平行关系.

因为这个群在所有的情况下起着重要的作用,因为它与坐标轴在通常空间中变换的群同构,还因为它如此密切地和我们的三维空间联系在一起,由于这些理由,当这个群以最自然的方式,即通过引入三维空间被提出时,我们的方程将取它们最简单的形式.

并且由于这个群本身与被认为固体的每一单元的位置变化的群同构,由于服从这个群的规律的运动固体的这一性质通过最终分析只不过是我刚刚注意到的不变性这一特征的特例,所以我们看到,在导致我们把三维赋予空间的物理学的根据和在前面提出的心理学的根据之间,并不存在基本的差别.

6 “拓扑学”和直觉

我想附加一点评论,它仅仅与我已经说过的东西间接有关.我们在上面看到了拓扑学的重要性,我解释道,在这里有几何学直觉的合法领域.这种直觉存在吗?我将回想起,存在着不要直觉也想取得进展的企图,而且希尔伯特先生试图建立一种所谓的理性几何学,因为这种几何学一点也不诉诸直觉.它以一定数目的公理或公设为基础,这些公理或公设被认为不是直觉的真理,而认为是伪装的定义.这些公理被分为五组.关于其中的四组,我已在某些场合提到了,在某种程度上把它们视为只包含伪装的定义是合理的.

在这里,我想着重强调一下其中的一组,即第二组,“次序公理”组.为了充分解释这个组涉及什么内容,我将引用它们中的一个.如果在任一线上的 A 和 B 之间有任意一点 C,在 A 和 C 之间有任一点 D,那么点 D 将处在 A 和 B 之间.按照希尔伯特先生的观点,其中没有直觉的真理;我们同意说,在某些情况下,C 在 A 和 B 之间,可是除了我们知道点或线是什么之外,我们不知道这意味着什么更多的东西.按照我们的法则,为了在任意三个点之间指定任何关系,我们能够使用“在……之间”这个表述,只要这个关系满足次序公理即可.于是,这些公理在我们看来好像是“在……之间”这个词的定义.

因此,有可能利用这些公理,只要满足这个条件,即证明它们不相互矛盾;而且,几何学也有可能建立在它们的基础上,在这种几何学中,将不需要图形,它能够被既没有视觉、触觉,也没有肌肉感觉以及任何感觉的人所理解,它可以





归结为纯粹的知性.

是的,这种人也许会在下述意义上来理解:他十分清楚地认识到,这些命题在逻辑上可以使一个从另一个中推导出来;但是,这些命题的集合对他来说似乎是人为的和奇异的,他不理解为什么是这种命题集合,而不是许多其他可能的集合更受欢迎.

如果我们没有经历同样的惊奇,那正是因为对于我们来说,公理实际上不是简单的定义和任意的约定,而是真正证明为正确的约定.至于其他各组公理,我依然认为,它们之所以被证明是正确的,是因为它们是与我们熟悉的某些经验事实最近似符合的东西,因而对于我们来说,它们是最方便的.谈到次序公理,在我看来,似乎存在着某种更多的东西:它们是与拓扑学有关的真实的直觉命题.我们看到,点 C 在一条线上其他两点之间的事实与借助于由不可逾越的点形成的截量去截取一维连续统的方法有关.

可是,接着便产生了一个问题:像次序公理这样一些真理是通过直觉向我们揭示出来的;但是,这是有关空间直觉本身的事情呢?还是有关一般的数学连续统或物理连续统直觉的事情呢?倘若赞成第一种解决办法,我们可以容易地论证空间,但是要论证更复杂的连续统、要论证不能在空间中来描述的大于三维的连续统就困难得多了.

而且,如果第一种解决办法被采纳,这里的全部讨论会变得毫无用处;我们之所以将三维性直率地赋予空间,是因为三维连续统是我们能够具有清晰直觉的唯一连续统.

但是,还存在着大于三维的拓扑学.我没有说它是一门容易的科学,我为此付出了巨大的努力,没有考虑到会在其中遇到这么多困难.但是,无论如何,这门科学是可能的,它并未全部停留在分析学上.要是不持续地诉诸直觉,就无法成功地把它探究下去.因此,确实存在着大于三维的连续统的直觉;与通常的几何学直觉相比,如果它要求比较持久的注意力,那么这无疑是一个习惯问题,也无疑是当维数增加时,连续统复杂性急剧增加的结果.我们难道在我们的中等学校没有看到平面几何学得很好的学生"无法想象空间"吗?那不是他们缺乏三维空间的直觉,而是他们不习惯于运用它,他们需要做出努力才能如此.而且,为了想象空间图形,我们难道不去相继地想象这个图形的各种可能的远景吗?

我将得出结论,我们大家都有任意维数的连续统的直觉概念,因为我们具有构造物理连续统和数学连续统的能力;而且,这种能力之所以在任何经验之前就在我们身上存在着,是因为没有它,经验严格说来是不可能的,会沦为不适合任何有机体的没有理性的感觉;是因为这种直觉只不过是我们具有这种本能

的意识.然而,这种本能可以以不同的方式来运用;它能够使我们像构造三维空间那样来构造四维空间.正是外部世界,正是经验,引导我们在一种意义而不是在另一种意义上运用它.





第十章

自庞加莱建立组合拓扑学开始,他以及早期的后继者都以三维流形作为主要的研究对象.这是自然的,因为二维的曲面分类已基本完成,三维的对象是可以借助直观作部分想象的,并且可依靠对曲面理论的利用以及对纽结理论的利用而获得利益.

组合方法得到充分发挥,基本群对于三维流形的重要性被认识.围绕着庞加莱猜想的努力却始终未获成功.但对于透镜空间,Seifert 流形和 Haken 流形等的研究成果有重要意义,为拓扑学提供了大量重要例子的库存.

重要的突破是于 20 世纪 70 年代由 W. Thurston 做出的.他将双曲几何结构引入三维流形分类而获得巨大进展.

下面就来介绍这一发展简史.注意到在三维情形,TOP 范畴、PL 范畴和 DIFF 范畴完全一致.因此,我们为了方便有时采用组合结构,有时采用光滑结构,又由于从可定向流形推广到不可定向流形并无根本性的困难,为简单起见,本章中谈到的三维流形都假定是可定向的.

庞加莱猜想

庞加莱于1900年在他的第二篇补充的最后,提出以下断言:同调平凡的流形同胚于一个球面.四年以后,他在第五个补充中举出一个反例,一个闭三维流形具有与 S^3 一样的同调,但其基本群不是平凡的.这个流形现今已被称为庞加莱流形,它的构作用到了 Heegaard 于1898年的博士论文中提出的办法,现通称为 Heegaard 分解.

设 M 是一个连通的三维流形.设 T 是 M 的一个单纯的三角剖分,使成为一个组合流形. T 的一维骨架在 M 中的一个正则邻域的闭包记作 H_1, H_1 在 M 中的余集的闭包记作 H_2,它可看做 T 的对偶三角剖分的一维骨架在 M 中的一个正则邻域的闭包.容易看出 H_1 和 H_2 都是在一个闭的三维胞腔上附贴上相同数目的一维环柄,即都是亏格为 g 的一维环柄体,它们享有公共的边缘,它是一个亏格为 g 的闭曲面.于是,(H_1, H_2) 称为 M 的一个 Heegaard 分解(splitting).

反之,设 H_1 和 H_2 是给定的两个亏格都是 g 的一维环柄体,并设 $h:\partial H_2\to\partial H_1$ 是一个同胚,则用 h 黏合后所得空间 $M=M_1\cup_h H_2$ 是一个三维流形,并且 (H_1,H_2)是 M 的一个 Heegaard 分解.

庞加莱当初的做法如下:取一个亏格为2的有向环柄体 H_1,设 $D_1^{(1)}$ 和 $D_2^{(1)}$ 是 H_1 中两个环柄的核心圆盘.将 H_1 沿 $D_1^{(1)}$ 和 $D_2^{(1)}$ 切开得一个三维球体,同时将 H_1 的表面∂H_1,一个亏格为2的有向曲面,切开成为一个有4个洞的球面,它可铺平为有3个洞的圆盘(图10.1).记圆盘的周界及3个洞的周界定向后分别为 $\pm A=\pm C_1^{(1)}$ 和 $\pm B=\pm C_2^{(1)}$,它们分别由 $D_1^{(1)}$ 和 $D_2^{(1)}$ 切开.另取一个亏格2的有向环柄体 H_2,设 $D_1^{(2)}$ 和 $D_2^{(2)}$ 是 H_2 中两个环柄的核心圆盘并且 $C_1^{(2)}=\partial D_1^{(2)}$ 和 $C_2^{(2)}=\partial D_2^{(2)}$ 是∂H_2 上的两个有向圆周.设 $h:\partial H_2-\partial H_1$ 是附贴同胚,则 $C'_1=h(C_1^{(2)})$和 $C'_2=h(C_2^{(2)})$称为特征曲线(Characteristic Curves).记 $M=H_1\cup_h H_2$,它被∂H_1 上的特征曲线 C'_1 和 C'_2 所决定至同胚.而且 M 的基本群可用如下方式表出:对 H_1 的每个环柄取一个生成元,然后在 $\pi_1(H_1)$中取特征曲线 C'_1 和 C'_2 代表的同伦类$[C'_1]$和$[C'_2]$为关系子,则关系为$[C'_1]=1$ 和$[C'_2]=1$.特别地,若取特征曲线如图10.1所示,所得三维流形即庞加莱所构成,今记作 P.于是有

$$\pi_1(P)=\langle a,b,;a^4b^{-1}a^{-1}b^{-1}=1,b^{-1}a^{-1}b^2a^{-1}=1\rangle$$

但 $\pi_1(P)=[\pi_1(P),\pi_1(P)]$,故 $H_1(P)=0$.再由庞加莱对偶定理,$H_2(P)=0$,即 P 是一个同调三维球面,但 $\pi_1(P)\neq\{1\}$.

庞加莱接着针对三维提出下面猜测.



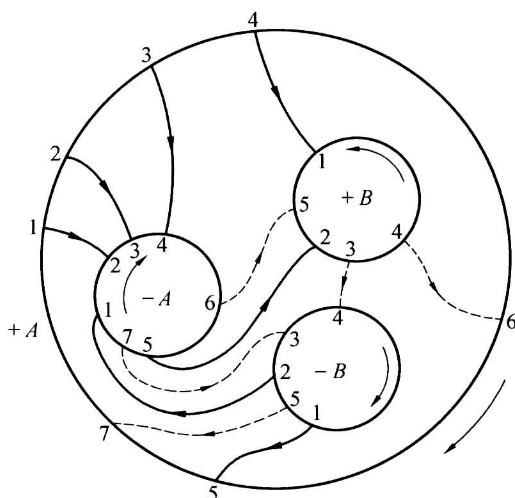


图 10.1

庞加莱猜想 每个单连通的闭三维流形同胚于 S^3.这里单连通意指连通并且基本群平凡.

同伦等价于 S^n 的闭 n 维流形称为同伦 n 维球面.不难证明,闭的三维流形是同伦三维球面的充要条件是,它是单连通的.于是庞加莱猜想可改写为

每个同伦三维球面均同胚于 S^3

将它推广到一般维数便有:

广义庞加莱猜想 每个同伦 n 维球面同胚于 S^n.

这是 TOP 范畴的提法,也可对 PL 范畴和 DIFF 范畴来提.

庞加莱猜想至今仍悬而未决.它在近百年的期间始终是数学家最感兴趣的问题之一.

直接为庞加莱猜想提供的一个"证明"是 1934 年 J·H·C·怀特海的论文,但这是错的.该作者自己接着发表了一个更正.1958 年一份日本杂志上的一篇长达一百零六页的文章也未被认可.1985 年报载英国瓦里克的数学家提供了一个解答,接着厚达八十页的预印本散发全世界,亦未被认可.至今,关于庞加莱猜想已获解决的传说时有所闻,从未间断.我们盼望从中会有一个得到数学界的认可.庞加莱猜想成立与否有重要的推论.例如,有下面的:

定理 若庞加莱猜想成立,则任何三角剖分的 4 维流形必为组合流形.

由此,从目前的知识可知,若庞加莱猜想成立,则存在闭的 4 维流形不可三角剖分,例如 $|E_8|$,因为它是不可光滑化的从而它不是组合流形.

三维空间里的拓扑等价关系[1]

第十一章

1 拓扑等价关系

一个立体球拓扑等价于一个立方体或任何正则立体.三维欧几里得空间中的两个图形称做拓扑等价的,当且仅当存在一弹性运动,使得其中一图形与另一图形重合.当然如果我们给出两个物理客体,其中一个是固体橡胶球,另一个是固体的木制立方体,我们不能使橡胶球与木立方体重合.如果我们想使它们重合,它们就得互相撞击,使得橡胶球平贴在立方体的外表面上而不通过立方体的内部.这一例子可以用来强调数学所研究的图形——无论是在欧几里得几何中还是在拓扑学中——不是物理客体,而是抽象图形.一个三角形不是可以用木头、纸张或绳子做成的东西,它是由以某种方法安放在一起的"线段"构成的,而线段又是"点"的某种组合.这样,对几何或拓扑进行任何认真的研究,就需要一个适当的基础知识,即对点集合进行一些讨论,并清楚地理解怎样使一个点集与另一个点集重合.现在,我们仍然依靠对图形的拓扑等价的直觉,它是建立在理想弹性图形的弹性运动上的.

下面将讨论我们感兴趣的几个标准图形,其中有的在前面的章节中非正式地出现过,但为完备起见,这里将它们描述一番.

① 原载自:[美]ARNOLD.初等拓扑的直观概念[M].王阿雄,译.北京:人民教育出版社,1980:52-64.



一个圆是一平面曲线,它上面的每一点与平面上一定点的距离为某一定长.这一定点就是圆心,定长即圆的半径.一单纯封闭曲线乃是与圆拓扑等价的曲线.一个单纯封闭曲线可以在一个平面上也可以不在一个平面上.图 11.1 给出不在一平面上的打结的单纯封闭曲线的例子.

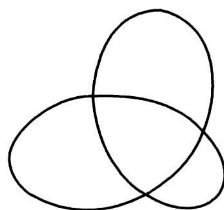


图 11.1

一个开圆盘乃是平面上被某一圆所围住的那部分,但不包括圆本身.一个闭圆盘乃是平面上的某一圆的内部或圆上的那部分,即圆曲线上的每一点都在闭圆盘上.注意,无论开的或闭的圆盘都是平面上的一个表面.一个闭的圆盘是一个开的圆盘加上其内部为该开圆盘的那个圆.

一个球面是一个三维的表面,它上面所有的点与某一定点的距离为某一定长,这个定点即此球面的球心,这个定长即此球面的半径.

一个开圆球是三维空间中被某一球面所围住的那部分,但不包括球面本身.一个闭圆球是三维空间中某一球面的内部或球面上的那部分(即球面上所有的点都在闭圆球上),注意,无论开圆球还是闭圆球都是三维空间里的一个立体.一个闭圆球是一个开圆球加上其内部为该开圆球的那个球面.

一个带 p 个柄的球面乃是三维空间里这样的一个表面:在一球面上凿 $2p$ 个洞,弯曲 p 个不同的管子,将它们的两端分别插在这些洞里而得到的一个表面.图 11.2 表示一个带三个柄的球面.

一个"胎形"(图 11.3)乃是三维空间里这样的一个表面:将一个圆绕圆所在平面且不与此圆相交的直线旋转而得到的一个表面.可以把胎形想成一个车轮的内胎或一个油炸圈饼的表面.

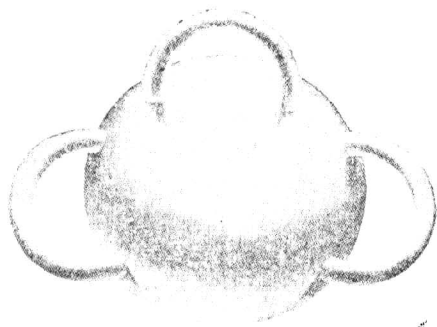


图 11.2

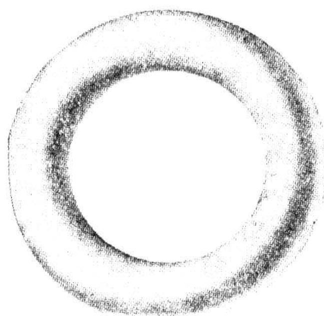


图 11.3

如果两个图形是拓扑等价的,我们可以用一弹性运动将一图形变成另一图形来证明之.例如,设有一圆形橡胶带,我们可将它切断,弯曲成图 11.1 中曲线的形状,然后再把两端接成它们原来的样子.这可以证明图 11.1 中的曲线确实

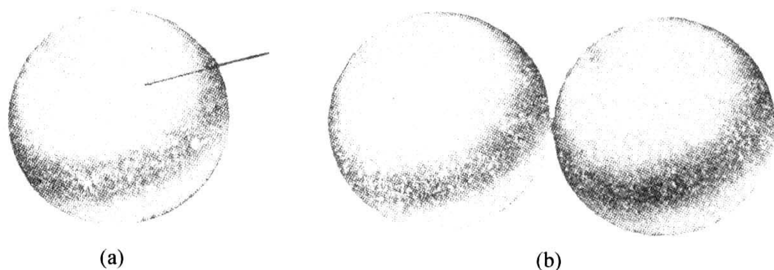
是一个单纯封闭曲线,因为它拓扑等价于一个圆.

那么怎样证明两个图形不是拓扑等价的呢?这就需要证明不存在任何弹性运动,可将一个图形变成与另一图形重合.当然,我们不可能一一试用每种弹性运动——不同的弹性运动太多了.给出这类证明的一个方法是,找出其中一个图形的一种性质,而这性质不为另一图形所具备.如果这一性质是拓扑性质,那么这两个图形就不可能是拓扑等价的.因为任何弹性运动都不能产生或消除这一性质,所以没有一种弹性运动可使其中的一个图形与另一图形重合.我们证明一个球面不拓扑等价于一个胎形,以此为例说明这类证明过程.事实上,如果球面上面有一任意单纯封闭曲线,并沿此曲线切割球面,则这个表面就分裂成两部分,故球面上任意单纯封闭曲线都使得这个球面不再是连通的,而胎形不具有这一性质.如果有一个圆穿过胎形的一个洞并围绕着胎的横切面,用这样的圆曲线切割此胎形,这个表面就变成一根管子,但它仍然是一个整体.所以胎形不能被这样的曲线切割成不连通的,故胎形不具有球面所具有的上述性质.而且,一个表面可以被它上面的任意单纯封闭曲线切割成不连通的这一性质显然是拓扑性质.于是,一个球面与一个胎形不拓扑等价.

2 表面的分类

每个学生都知道一张普通的纸有两个面,有时他必须只在一面上写字,有时则可在两面上都写,在这一节中,我们将看到一些只有一个面的表面的例子,但首先必须稍微详细一点描述,我们所要考虑的到底是哪些表面.

像平面或球面一样,一个表面可以是“两维的”;但是,伸出一根刺的球面(图 11.4(a))或两个相切的球面(图 11.4(b))的情况怎样呢?它们是表面吗?本节我们考虑一种叫做流形的特殊表面.一个流形是一个连通的表面(即一个“一整片”的表面),其中,极其靠近表面上任何一点的局部表面都拓扑等价于一开圆盘.这就是说,对表面上的每一点 p,此表面上所有与点 p 足够近的点构成一个与开圆盘拓扑等价的集合,这个由在表面上所有靠近点 p 的点所构成的集合叫做 p 的邻域.



(a) (b)

图 11.4

图 11.4 中所示表面都不是流形,在图 11.4(a)中,在刺上的点没有满足所需条件的邻域.在图 11.4(b)中的表面上在接近两球面切点的地方不与圆盘拓扑等价.

一个表面是“有界”的,当且仅当整个表面都包含在某个开圆球中.胎形是有界表面,平面不是有界表面.如果考虑一个表面的某一片特定部分时,这片表面的“边界”定义为把这一片与表面的其余部分分开的那条曲线.例如,把圆盘视为平面的一部分,圆盘的边界就是包围它的那个圆;平面上一个圆环的边界包括两个圆.注意,表面上的某一片的边界不完全取决于这一片本身,而且也取决于表面的其余部分,因为边界必须是两个集合的“划分者”或“边界”.这一点可以用一个例子说清楚.如上所说,若将圆盘视为平面上的一部分,它的边界就是一个圆.假如我们把一个圆盘作为该圆盘本身的一部分来考虑,则表面上所有的点都在所考虑的这一部分里,于是当然不存在这一部分和表面上其余部分之间的“划分者”.即,当把圆盘看做是它自身的子集时,它没有边界.现在,我们只关心一些简单的情况,所以只要用直觉概念就够了.当我们说到一个表面的边界时,我们总是认为这已知表面是其自身自然延伸的那个表面的一部分,例如平面里的圆盘.注意到下面一点也是很重要的.即尽管“有界的”(bounded)和“边界”(boundary)这两个词很相像,这两个概念却是根本不同的.一个球面和一个圆盘都是有界的表面,但是球面没有边界,而圆盘的边界是一个圆.一个平面是没有边界的无界表面;平面上的一细条是有两条平行线为边界的无界部分.一个表面是“封闭的”,当且仅当它是有界的且没有边界.一个球面是一个封闭的表面,因为它是有界的,并且我们可以在一个球面上自由运动而永远不会碰到一个边.另一方面,无论开的还是闭的圆盘都不是封闭的表面,因为它们都有包围着它们的圆作为边界.在这两种情况下,它都是指运动不会为一个端点或一条边所阻止.(一个封闭的路径不要求有界性这一条件,因为所有路径都是有界的),一个流形可以是也可以不是一个封闭的表面.球面是一个流形,它是封闭的;开圆盘也是一个流形,但它不是封闭的表面.闭圆盘不是一个流形,因为在边界圆上一点的靠近地方,其表面不拓扑等价于开圆盘.

有一个有趣的方法来表示某些流形,即把长方形的某些边黏合起来.例如,图 11.5 中,把长方形的两端黏合在一起,使得标记为 AB 的两条线段重合,并使图中的两个箭头方向一致,这个长方形所表示的流形就是一圆柱曲面.注意线段 BB,AA 代表圆柱表面两端的圆.在这两个线段(或圆)上的点不能包括在流形上,因为它们在表面上没有合适的邻域.圆柱的这种表示法有时很方便,

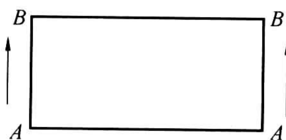
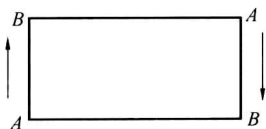


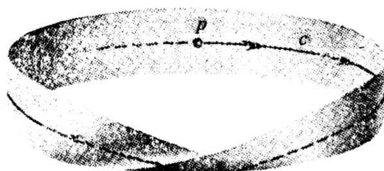
图 11.5

因为它能平放在桌上而且立即可以看到表面的一个面的全部.

假设图 11.5 中的一个箭头倒过来,如图 11.6(a)所示.如果把长方形的一端扭 180°,然后将两端黏合在一起,使得标记为 *AB* 的两个线段重合,并且使两个箭头的方向一致,就得到一个叫麦比乌斯(Möbius)带的流形(图 11.6(b)).为了得到一个流形,必须像图 11.5 中那样,这个表面必须不包括图 11.6 中的两条水平线段.然而这两个线段不再代表两个圆了,这里它们代表同一单纯封闭曲线的两半.一个麦比乌斯带只有一个面(Side),为了了解这一点,设想一只苍蝇从点 *P* 出发沿表面上的曲线 *C* 爬行,它沿着 *C* 的虚线部分回到点 *P*.由于苍蝇既没有穿过表面,也没有绕过表面的边,所以它一定一直在表面的同一面,但是看起来,它是从表面上它所出发的"另一面"回到点 *P* 的.因此看起来是点 *P* 所在表面的两个不同的面,实际上只是此表面的同一面的两个部分.注意单面性(Onesideness)与其说是表面的内在属性,倒不如说是它的外在的属性.这就是说,检验一个表面是否是单面的,不完全是在表面上进行的,而是使用了周围的空间.在上面苍蝇的例子中,我们已经默认:如果苍蝇走动时脚一直在表面上,并且不通过表面的任何边,则它的头的运动所经过的所有的点也一定在表面的同一面.如果表面是一麦比乌斯带,这只苍蝇的头就能从任意接近表面的点运动到任何另一接近表面的点,于是所有点都在表面的同一面,这表面是单面的.



(a)



(b)

图 11.6 麦比乌斯带

图 11.7 给出一个面的表面的另一个例子.这个流形叫做克莱因(Klein)瓶,它不可能在三维空间中造出来.图 11.7(a)给出克莱因瓶的表示:一个两对边分别黏合的长方形.用实箭头标出两条边的黏合给出一个圆柱形表面(图 11.7(b)).它的两端还有待重叠在一起.这一步骤不可能在三维空间中进行(图 11.7(c)).圆筒的一端一定要戳过面而从里面与另一端相接.但是,这个表面在瓶颈通过面的地方不能与其自身相交(这就需要一个四维空间来使得瓶颈"绕过"这个表面而不是"穿过"它).

下面的定理给出一个有趣的对封闭的双面流形的分类法,类似的对单面流形的分类法可在所给的参考文献中找到.

定理 任一封闭的双面流形均拓扑等价于一个带有几个柄的球面.



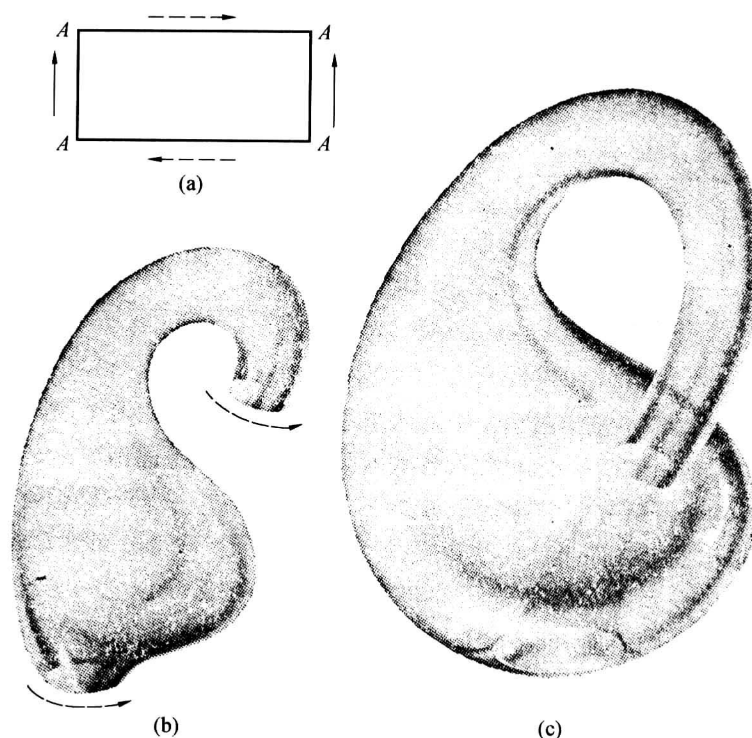


图 11.7　克莱因瓶

证明　本定理的证明超出本书的范围.有兴趣的读者可以再参考 Brumfiel, Eicholz, Shanks 的书《Geometry》33 页定理 2 或参考 Aleksandrov 的书《Combinatorial Topoloby》vol. 1 第 110 页定理 7.2.

什么是拓扑学[①]

第十二章

1 克莱因的定义

克莱因在 1893 年发表的爱尔兰根计划书(Erlanger Program)中,对几何学及其各分科下了定义和分类,一度促进了几何学的发展.他的想法基调是简单的.

例如,欧氏几何是研究图形(作为刚体)在运动中不变性质的.运动的全体形成一个群(Group),叫运动群(凡是使某种性质不变的空间的变换构成一个群).因此,克莱因就把欧氏几何看做运动群的不变量论.射影几何、仿射几何、各种非欧几何等都分别是射影变换群、仿射变换群、非欧运动群等的不变量论.各种几何与变换群互为从属(Subordinate).拓扑学是空间(在克莱因那里意味着欧氏空间)的从属于最一般的一对一连续变换群的几何——这就是克莱因原先的定义.

n 维实数空间 $\mathbf{R}^n$ 的一对一连续变换称 $\mathbf{R}^n$ 的拓扑变换.简写为

$$x \rightarrow x'$$

当 $\mathbf{R}^n$ 里的两个图形之一按 $\mathbf{R}^n$ 的拓扑变换可以变到另一个时,我们说:这两图形在 $\mathbf{R}^n$ 里是同痕的.根据这些定义,我们可以这样说:$\mathbf{R}^n$ 空间的拓扑学是研究图形对于 $\mathbf{R}^n$ 的拓扑变换不变性质的几何.

① 原载自:苏步青.拓扑学初步[M].上海:复旦大学出版社,1986:58-78.



例如,在 $\mathbf{R}^2$ 的场合考察一只橡皮面,并且想象伸缩(既不撕破、又不折叠)之后的平面,我们便会领会到"$\mathbf{R}^2$ 的拓扑变换"的直观印象.在橡皮面上,比方画一个圆,经过伸缩,就得出各式各样和圆"在 $\mathbf{R}^2$ 上同痕的"曲线(例如图 12.1(a)),这些曲线具有的共同性质——此时,这些曲线把面划分为内外两部分,等等——就是所谓拓扑的性质.值得注意的是,图 12.1(b)在三维空间里和圆是同痕的,但是和图 12.1(c)则不同痕.

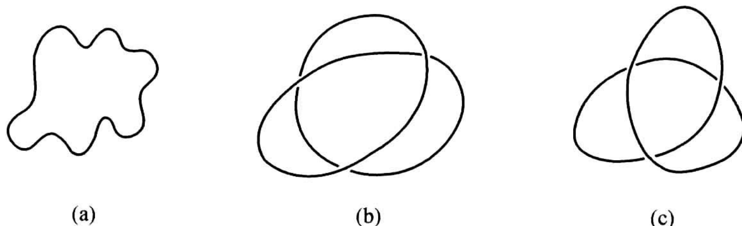


图 12.1

由于变换群本身是按代数化被定义的,克莱因相信:用上述定义便可完全将几何和其各分科代数化(或抽象化).可是对一般的黎曼几何等却无法如此给以特征化(Schouten, 1924),这一度使克莱因感到有点狼狈①.关于拓扑学也是如此,限于上述定义来考虑问题,会束缚人们的思想.对此,将在下一节叙述.

2 位置与拓扑

如前节所述,定义的拓扑学单是研究欧氏空间 $\mathbf{R}^n$ 内的图形 F 对于 $\mathbf{R}^n$ 的相对位置.但是,有时还遇到图形 F 本身的内在构造的问题,就是与周围空间无关的问题.

例如,图 12.2 中的两图形在平面上并非同痕,但在三维空间里却是同痕的(图 12.1(c)也有同样的情况,它在四维空间里是和圆同痕的).两图形有相同的内在结构.

从一个图形 F_1 到另一个图形 F_2 的一对一双方连续映射(Map)称拓扑映射.从 F_1 经过拓扑映射可以得出 F_2 时,我们说 F_1, F_2 是同胚的(Homeomorphic).研究同胚图形的共同性质,就是拓扑学的重要课题.

关于同痕、同胚便分别产生两个主要问题:同痕问题,就是在给定的空间里

① 凡可用克莱因定义解释的几何所在的空间称克莱因空间.E·嘉当非常巧妙地把一般黎曼空间看为处处由克莱因空间所构成而且加以欧氏联络.

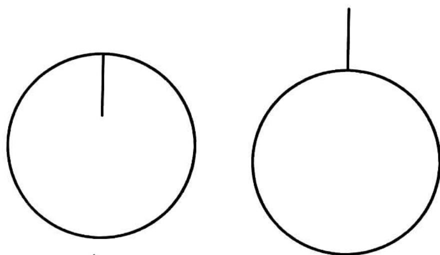


图 12.2

决定两图形同痕的充要条件的问题；同胚问题，就是求两图形同胚的充要条件的问题.前一问题比后一问题更为困难.三维空间的闭曲线(像图 12.1(b)、(c)一类的曲线)在什么场合下才和圆同痕，这个所谓的“打结问题”(Knots problem)是其中最简单的特殊情况，但是到如今还不能说已得到解决的头绪①.至于同胚问题，如下节所述，对曲面已经解决了.

另外，在克莱因的定义中，$\mathbf{R}^n$ 常被看为欧氏空间，那未免太狭隘了.“空间”这个概念应该大大扩充，这里姑且不谈，把 $\mathbf{R}^n$ 看做为任意的图形 F_1，F_2 被画于其上的图形(比方，$\mathbf{R}^n$ 是球面，F_1，F_2 是球面上的曲线)，而假定按照 $\mathbf{R}^n$ 到它本身去的拓扑映射可将 F_1 移到 F_2.这时，我们可以说：F_1，F_2“在 $\mathbf{R}^n$ 里是同痕的”.这样的相对关系当然在拓扑学里该成为讨论的对象，并且拓扑映射本身或者(不一定是拓扑的)连续映射也该是研究的目标.总之，拓扑学的对象并非是按照克莱因的定义限制起来的东西，而大体上把它作为连续的几何学(Alexandroff-Hopf 著《拓扑学》的前言中的用语)，该是得了要领的说法.

3 曲面的同胚问题

黎曼在 1851 年奠定了函数论的基础，首先提出“面”(或称二维流形)作为研究的对象.学过椭圆函数论的读者都知道，代数函数

$$y^2 = a_0x^4 + a_1x^3 + \cdots + a_4$$

的黎曼面和环面是同胚的.对曲面同胚问题的系统研究就是从他开始的，而且由麦比乌斯(1863)、若尔当(1856)、雪莱夫里(1872)、台克(1888)等人加以解决.就是说，曲面的拓扑决定于下列三个拓扑学的元素.

① 关于打结问题可参考 Reidemeister 的著作 *Knotentheorie*.





(1)境界的数目.在球面、环面,是0;在圆板,是1.对闭曲面用针穿孔,有几个孔,就有几个境界.一般地,它的含义是显而易见的.

(2)欧拉的示性数.在面上引任意条线把它"分割"成为像多面体的图形.这时,设"顶点"的个数为 a_0,"棱"的条数为 a_1,"面"的张数为 a_2,那么可以证明:$a_0 - a_1 + a_2$ 这一数和"分割"的方法无关,从而是一定的.(在球面,它等于2;在环面,它等于0.那里的证明可用于一般情况,望读者自证之.)这个数称欧拉的示性数(Characteristic).

(3)可定向性(Orientability).球面、环面等所谓"普通的"面是可定向的,或者说,是有表有里的.它的含义如下述.在面上一点的周围画一个可定向的小圆(或者从面的一侧指向他侧,即从"里"到"表"引小法线),而且把这小圈(或者法线)和这点一起在面上自由而连续地移动,使它回到原位置时,圆(或法线)的方向一定与出发时的方向一致.换句话说,可以区别表里.麦比乌斯最初指出,存在有不具备这性质的曲面.图12.3示意了不可定向的最简例子,即所谓"麦比乌斯带"(把细长的矩形纸扭一次而且贴合其两端,便可获得).

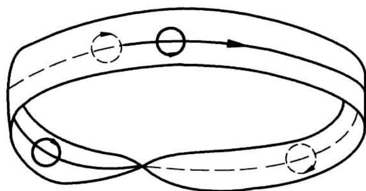


图12.3

这里,境界的数目和欧拉示性数分别相等,而且同是可定向或者不可定向的三个性质是两个面同胚的充要条件.证明见于拓扑学书籍,此处从略.

在可定向的面上,设 r 是境界数,$a_0 - a_1 + a_2 + r$ 一定是偶数.把它写作 $2(1-p)$ 时,p 称面的亏格(Genus, Deficiency).p 永远大于等于0,在球面是0,在环面是1.这个数在代数函数论里具有重要的意义.(凡可表成黎曼面的面,只有可定向的面)

4 近百年来发展的两个方向、基本群

我们在上述的章节中,主要根据直观考虑问题,而没有对"曲线"、"曲面"等词汇深入研究其含义.从数学上看,这未免是不精确的了.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

19世纪后半期是数学的大检查时代,处处没有切线的连续曲线(魏尔斯特拉斯,1870年前后)、塞满一个正方形的连续曲线(皮亚诺,1894年)等的出现,要求我们正确地定义"曲线"、"曲面"等,这是当然的.此外,以往被认为"自明"的事项,现在变为不证明就靠不住的了.若尔当在其1893年出版的数学分析教程(Cours d'Analyse)中,以分析的严密基础化为目标,对曲线作为线段的连续象给了定义①,而且对圆的拓扑映象(称若尔当曲线)证明它把平面平分为内外两部分.

皮亚诺(Peano, Giuseppe, 1858—1932),意大利数学家、逻辑学家.

这个时代也可以说是"数学抽象化"的萌芽时期.就几何学说来,n维几何学早已为格拉斯曼(从1840年代)所创立,即使在三维空间里的事情,如果把它作为n维的特殊情况处理,反而可以透彻些.高斯、波尔约、罗巴切夫斯基等创造了非欧几何,从而促进了几何学发展.黎曼1854年的就职演讲(Habilitationsvortag)叙述了抽象空间的构成思想.如上述,克莱因的爱尔兰根计划书也是这个时代精神的一种表现.

格拉斯曼(Grassmann, Hermann Gunther, 1809—1877),德国数学家.

波尔约(Bolyai, János, 1802—1860),匈牙利数学家.

罗巴切夫斯基(Побачевский, Ннкопай иванович, 1792—1856),俄国数学家.

在这样的气氛中,经过两位数学家的努力,拓扑学在两个不同的方向上得到了飞跃的发展,这两位就是庞加莱和康托.

现在,先就庞加莱从1881年开始的一系列业绩作一般性的介绍.其中对于拓扑学全貌有巨大影响的结果形成了组合拓扑学(Combinatoric Topology)的科学基础.

庞加莱首先注意到把拓扑学的对象限制在三维"直观的"空间之内去找寻,在数学上是没有意义的.历史上,考虑n维拓扑学的最初数学家不是庞加莱而是Schäffli(1852),后者把欧拉公式拓展到n维空间.就是:

在"n球面"(n-sphere:和$(n+1)$维空间的超球面同胚的图形)的分割中,设a_k是$k(k=0,1,2,\cdots,n)$维元素的个数,那么成立下列公式

$$a_0-a_1+a_2-\cdots+(-1)^n a_n=1+(-1)^n$$

比方说,我们还可更一般地考虑n维空间里的m维流形(Manifold).例如,由n个变数$x_1,x_2,\cdots,x_n$的$n-m$个独立的解析函数$f_k(x_1,x_2,\cdots,x_n)(k=1,2,\cdots,n-m)$等于0所表达的$m$维流形(其中函数行列式不等于0),是可以作为拓扑研究的对象,而且这项研究从数学上来看,比起三维空间的游戏问题重要得多.庞加莱于是着眼于一般维流形的拓扑学,而且为了严格打好基础,他发现,把流形作为"曲多面体"(Curved Polyhedron),也就是和多面体同胚的图形来处理,较为方便.这就是组合法的出发点.庞加莱就这样对流形,更进一步地对任何曲多面体导进两个拓扑不变(Topologically Invariant,即同胚图形共有的)群,

① 从上面提到的皮亚诺曲线可以塞满正方形的情况看,这个定义不是确切的.后来E·嘉当给了另外的定义(见后文).



就是基本群(Goundamental Group)和贝蒂群(贝蒂 Group)①.

基本群或称"路群"或称"同伦类群"(Homotopy-class Group),举例来说明一下.环面 F(图 12.4)上的"路"是指在 F 上所画的可定向的闭曲线.两条路在 F 是同伦,指的是根据 F 上的"连续变形"(Deformation)可以把一方移到另一方.这时以记号≈表达它.例如,在图 12.4 中

$$a \approx a', b \approx b', c \approx c'$$

在变形过程中,容许自己和自己相交,所以 $c \approx c'$,但是,$a \not\approx a''$,$a \not\approx b$.凡是同伦的路都被看成属于一"类",而不加区别.上述的 a, a' 属于同一类,而 b, b' 则属于另一类.先走一类的路,接着走别类的路,其结果称两类的"结合"(Composition).这样一来,路类在这个结合下构成一个群,这就是 F 的基本群.

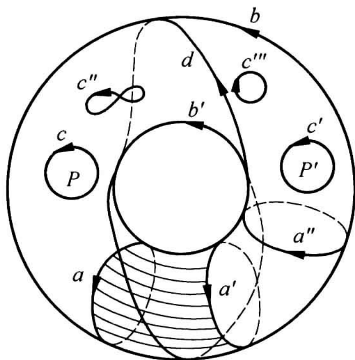


图 12.4

我们通常对两个路 a 和 b 的结合(先 a 后 b)写为积 ab.那么在图 12.4 上容易看出

$$a^{-1} \approx a'', c^{-1} \approx c''', d \approx a^2 b$$

5 贝蒂群

贝蒂群也称"同调类群"(Homology-class Group),就 n 维曲多面体来说,有着从 0 维到 n 维的各种贝蒂群.为说明一维贝蒂群,我们仍采用环面作例子.这里成讨论对象的是 F 的闭曲线.这样两条闭曲线在 F 上同调(Homological),意味着:把其中一条的方向反过来而得到的闭曲线和另一条恰恰构成 F 上的(有向)曲面的(可定向)境界.上述的同伦路也是同调的,但是在图 12.4 中的 P 和 P' 处,把 F 穿孔使成为新的曲面 F',那么 c 和 c' 关于 F' 就不是同伦的,但它们还是同调的.现在,在同一图里,改变 a' 的方向,使 $-a'$ 和 a 构成一个曲面的境

① 贝蒂于 1870 年导入此群,庞加莱给了严密的定义.

界，图 12.4 中附有阴影线的就是这个曲面可定向的面(沿境界正方向前进时，面常在它的一侧(左侧)，参见图 12.5)．那么，按上述定义 a 和 a' 是同调的，并且也是同伦的．同调的闭曲线属于一个类，于是我们把前面的同伦换作同调而用为结合，同样得到群，就是一维的贝蒂群．

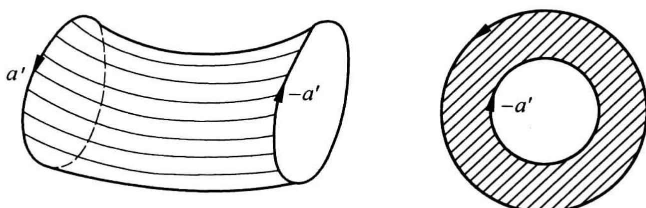


图 12.5

在环面上，一维贝蒂群恰恰与基本群是同构的(Isomorphic)，但一般则不成立．比如，在上述的 F' 上已经看到这一情况．贝蒂群和基本群不同之处还在于：贝蒂群是阿贝尔群(Abelian Group)，所以它的结合是以加法来表达的．设 ~ 表示同调，那么关于 F 或 F' 成立下列关系

$$a \sim a' \sim -a'', b \sim b', c \sim c', d \sim 2a + b$$

这个同调性和贝蒂群的概念形成了现代拓扑学的中心思想，可惜的是在这里没有充分说明的机会．读者可参考一本很得要领的著书：

P. L. Alexandroff: Einfachset Grundbegriffe der Topologie(拓扑学最简单的一些基本概念)．我们仅仅就这个概念如何在"组合论"中被定义起来的大致情况讲解如下．

为此，首先讲几个定义．第一，单形．n 维欧氏空间的点、线段、三角形、四面体以及一般维的扩充，比如 m 维图形：设 $\boldsymbol{x}_0, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_m$ 为 m 维欧氏空间里但不在 $(m-1)$ 维线性空间里的 $m+1$ 个点．做出点

$$\lambda_0\boldsymbol{x}_0 + \lambda_1\boldsymbol{x}_1 + \cdots + \lambda_m\boldsymbol{x}_m$$

其中

$$\lambda_k \geqslant 0, \sum_{k=0}^{m}\lambda_k = 1$$

这样点集称 m 维单形(Simplex)．那么多面体可以作为沿各边连接着的有限个或可数个单形的和集来定义．构成多面体的这些单形和它们的边一起构成所谓(几何)复形(Complex)．按照单形的顶点排列便可对它"定向"．如果对它们的排列施行奇置换(Odd Substitution)，我们说：这个单形的定向变为"反定向"．对一个复形 K 所含的单形添上定向，而记作为 x_i；对其反定向记作 $-x_i$，然后给 x_i 附上任意的有理整系数 t_i．这时，称

$$C = \sum t_i \cdot x_i$$

为 K 所含的代数复形(Algebraic Complex).由于代数复形是可按一次型的算法作加减的,而关于加法构成了阿贝尔群,当然是无疑的.如果 K 所含单形的个数是有限的,那么这个群是有限基的所谓"自由阿贝尔群",把它记作 $L(K)$.

其次,边界.设 r 维单形 x 决定于 $(r+1)$ 个顶点 $a_0, a_1, \cdots, a_{r+1}$,即

$$x = (a_0, a_1, \cdots, a_{r+1})$$

我们定义下列的代数复形 ∂x 为 x 的边界(Boundary),即

$$\partial x = (a_1 a_2 \cdots a_{r+1}) - (a_0 a_2 \cdots a_{r+1}) + (a_0 a_1 a_3 + a_{r+1}) + \cdots +$$

$$(-1)^r (a_0 a_1 \cdots a_r) = \sum_{i=0}^{r} (-1)^i (a_0 a_1 \cdots a_{i-1} a_{i+1} \cdots a_{r+1})$$

对于一般的代数复形

$$C = \sum t_i x_i$$

定义边界为

$$\partial C = \sum t_i \partial x_i$$

从

$$\partial(C_1 + C_2) = \partial C_1 + \partial C_2$$

得知,凡边界为0的复形全体构成了 $L(K)$ 的子群.这样的复形称循环(cycle ≡ zyklus),这个子群记作 $Z(K)$.

另一方面,我们容易证明

$$\partial(\partial C) = 0$$

所以,凡是作为一个复形的边界的复形是循环,其全体构成了 $Z(K)$ 的子群,记作 $R(K)$. K 所含单形中有个最大维数的单形,这个维数称复形 K 的维(Dimension).设 K 是 n 维复形,任何整数满足 $0 \leqslant r \leqslant n$ 时,仅由 K 所含的 r 维单形 x_i^r 全体组成的"同维代数复形" $\sum t_i x_i^r$ 又构成 $L(K)$ 的子群,记作 $L^r(K)$.如同对 $L(K)$ 那样,从 $L^r(K)$ 导出 $Z^r(K)$, $R^r(K)$ 等群.商群

$$Z^r(K)/R^r(K)$$

就是 K 的 r 维贝蒂群.

这个群可以不是有限自由阿贝尔群,而且有时具有0以外的有限次的元,从而分解为"有限部分"(即由所有的有限次元组成的非0子群)和"自由部分".这个有限部分称 r 维挠群(Torsion Group).自由部分的基数称 r 维的贝蒂数(Betti Number),记作 $p^r(K)$.贝蒂群决定于挠群和贝蒂数.

以上所述,全是有关几何复形的议论.对多面体 P 分割为单形使变为复形,方法可以有无数种,因此,上述复形有关的内容表面上看来,和多面体 P 无关.可是到了1915年,J. W. Alexander 证明了,前面定义的贝蒂群仅由 P 得到决定,和分割无关,而且同胚的多面体产生同构的贝蒂群.这样一来,多面体本

身便决定了贝蒂群.更一般地,对曲多面体也同样成立.

上面所述的定义形式上都是非常抽象的东西,可能不易理解.读者可就环面证实这个定义和前面说明过的事实相符合.在环面的场合,挠群 = 0,$p^0 = 1$,$p^1 = 2$,$p^2 = 1$.

庞加莱把"n 维流形"看做连通曲多面体而且各点具有与 n 维球同胚的邻域,曲多面体是相当一般的点集,我们前述的各个问题都可被解释为曲多面体的图形,因而数学上就这样确立了令人满意的观点.此外,在两个曲多面体里,各维贝蒂群的一致是这两个曲多面体同胚的必要条件.这虽然不是充分条件,但是一方面,因为贝蒂群是阿贝尔群,因而有代数上容易处理的优点,而且另一方面,这些群决定了图形的很多拓扑性质.这里举两个例子.一个是闭曲面 F 的定向可能或不可能问题决定于 $p^2(P) = 0$ 或 1;另一个是曲面的欧拉示性数

$$a_0 - a_1 + a_2 = p^0 - p^1 + p^2$$

环面上的欧拉公式作为特殊情况,也可包括在这一结果中.更一般地,在 n 维多面体上,设 k 维单形的个数为 a_k,那么成立下列的欧拉 - 庞加莱公式

$$a_0 - a_1 + a_2 - \cdots + (-1)^n a_n = p^0 - p^1 + p^2 - \cdots + (-1)^n p^n$$

这些事实对于庞加莱的方法赋予了特异的价值.按照这个方法创出的所谓"同调拓扑学"(Homology-topology),在 20 世纪 20 年代里,由于 Alexander、Vebien、Lefschetz 等人的研究而大大发展了.

读者可参阅两本名著:

O. Veblen: Analysis Situs(1922).

S. Lefschetz: Topology (1930).

6 康托的集合论

康托在 1870 年前后创造了抽象集合论和点集论.在欧氏空间里,除了曲多面体等外,他观察了"任意的"点集,从而成了拓扑学的另一方面的开拓人.尽管曲多面体是相当"一般的"点集,人们并没有理由排除对其他点集研究的想法."数学的本质在于自由",这就是康托的一贯主张.在他的点集论中有一些最基本的概念如凝聚点及其伴随的内点、外点、边界、闭集、开集等,都是拓扑学的概念,即经过拓扑变换不变的概念.今天的拓扑学中,这些术语是非有不可的了!康托还导进连通的概念,其定义是这样的:设一个集合不可能分成两个既不是空集又不具有公共点的开集,那么称这样的集合是连通的.从此引申出连续统(Continuum)的概念,定义:包括两个以上的点而且是连通的闭集.他还提示了对平面上的曲线的一个定义(不具内点的有界连续统),即所谓康托曲线的定义.





这门点集拓扑学(Set-theoretical Topology)为后来勒贝格的积分论、侧度论(Measure Theory)的发展所刺激,结果是:一派形成了所谓"画法集合论"(descriptive theory of sets; Lusin、Suslin 等人的所谓 Borel 集合论、解析集合论、射影集合论),而另一派成了以亚尼谢夫斯基为首的"波兰学派",1910 年以后研究了所谓"局部连通的连续统".①特别是乌利松研究了康托曲线的形状,而且闵戛对一般点集的维数理论试作了集成工作,1928 年写出了一部庞大著作 *Dimensionstheorie*.后来,他得到了纳柏林的协助,彻底研究了一般的一维图形(曲线).参照 K. Menger: Kurventheorie(1932).

7 一般拓扑学

康托考察了最一般的点集,但他只采取欧氏空间里的点集.对于拓扑学的成立,这样做不一定是必要的.不但如此,在数学分析中的应用要求我们研究更加一般化或"抽象化"空间,即所谓抽象空间的拓扑学,这是费雷歇(1906)最早注意到的.

费雷歇从泛函分析(Functional Analysis)着手研究.如所知,古典的数学分析中的"函数"是使一个(或 n 个)数对应于其他一个数的.但是,把一个"函数"看做"变数",而且考察同它对应的数或函数,这桩事在数学分析中是屡见不鲜的.例如,对于一个在定区间里定义的函数求在这区间上的定积分的数值,或者求不定积分,使与之对应,等等.这时,作为"变数"看待的"函数"称为"点",而且它的集合称为函数空间.当然,这样的空间不能表成为有限维欧氏空间,但在某种程度上存在着类似.关于这种"空间"的拓扑研究,应用上尤为重要②.

费雷歇(Fréchet, Maurice-René, 1878—1973),法国数学家.

要成立拓扑学,主要是定义连续映射.大体上说,把邻域映射到邻域的映射称连续映射.所以只需定义"邻域"的概念就可以.为此,先定义拓扑空间③:

设 R 为具有某些性质的元素的集合.假如对集合 R 的一个子集 U 都对应有 R 的子集 $\overline{U}$(称为 U 的闭包),使满足下列条件:

(1)假如 U 只含一个元素 x,那么 $\overline{U}=U$,即 $\bar{x}=x$.

(2)假如 U 和 V 是 R 的两个子集,那么 $\overline{U\cup V}=\overline{U}\cup\overline{V}$,就是说,和的闭包等于闭包的和.

(3)$\overline{\overline{U}}=\overline{U}$,就是说,两次运用闭包运算,其结果等于运用一次.

这样,我们就称集合 R 为拓扑空间.

其次,让我们来定义邻域.

① 杂志 *Fundamenta Mathematica* 就是这个学派创办的.
② M. Fréchet: Espaces abstraits, Borel 丛书.
③ 邦德列雅金著.连续群(上册).曹锡华,译.1957:55-56.

如果 $U=\bar{U}$，称 U 是闭集，$R-U$ 称开集. 一个集合 $U(x)$ 称为点 x 的邻域，如果有一个开集 U，使得 $x\in U\subset U(x)$. 所有的 x 的邻域 $U(x)$ 形成一邻域系 $U(x)$，它有下列性质：

(1) $x\in U(x)$.

(2)如果 $U(x)$ 和 $U'(x)$ 是同一点 x 的两个邻域，那么必有 x 的邻域 $U''(x)$ 使得 $U''(x)\subset U(x)$，$U''(x)\subset U'(x)$.

(3)如果 $y\in U(x)$，必有 $U(y)\subset U(x)$.

除了上述的三条性质外，如果对拓扑空间 R 的邻域加上第四条，即

(4)对于两不同点 x 和 y 必存在 $U(x)$ 和 $U(y)$ 使得 $U(x)\cap U(y)=\varnothing$，就是两邻域没有公共点，那么称这样的拓扑空间为豪斯道夫空间①. 公理(4)也称豪斯道夫公理. 这种拓扑空间特别重要. 它的一种特殊情况是所谓距离空间，定义如下：

设 D 是由元素(称点) $x,y,z,\cdots$ 组成的集，对于 D 的两元 x,y 有对应的非负实数 $\rho(x,y)$，称"距离"，使满足下述的三个公理：

(1) $\rho(x,y)$ 当且仅当 $x=y$ 时等于 0.

(2) $\rho(x,y)=\rho(y,x)$.

(3) $\rho(x,y)+\rho(y,z)\geqslant\rho(x,z)$.(三角形公理)

称这样的集合 D 为距离空间.

对于 D 的一点 x 取定邻域

$$y\in U(x)\,(\rho(x,y)<\varepsilon)$$

便可明白：距离空间是豪斯道夫空间.

上述的函数空间也可以作为距离空间来考虑.

对于豪斯道夫空间在什么场合下会变为距离空间的问题，乌利松解决了. 欧氏空间、希尔伯特空间当然都是距离空间.

现代所谓 n 维流形是指它的每一个邻域都和 n 维空间球(连同球体)同胚的拓扑空间. 此外，空间"度量化"(Metrization)就是：指定满足上述三个公理的"距离" $\rho(x,y)$ 与两点 x,y 对应. 所谓空间"拓扑化"是指"邻域系"的导进. 这样的抽象空间里的点集拓扑学称一般拓扑学(General Topology)，波兰数学家如库拉托夫斯基、希尔宾斯基等做出了许多工作. 我国学者江泽涵、吴文俊、张素诚、姜伯驹等，在拓扑学的领域里做了很好工作.

读者进一步可参阅：江泽涵. 布劳威尔拓扑学引论[M]. 上海：上海科技出版社，1978.

豪斯道夫(Hausdorff, Felix, 1868—1942)，德国数学家.

库拉托夫斯基(Karatowski, kazimierz, 1896—1980)，波兰数学家.

希尔宾斯基(Sierpinski, Wactaw, 1882—1969)，波兰数学家.

① F. Hausdorff: Grundlage der Mengenlehre (1914).





8 布劳威尔

荷兰数学家布劳威尔首先把组合方法应用于点集拓扑学,在 1911 年发表的关于维数不变性的论文成了拓扑学新时期的开端.

布劳威尔(Brouwer, Luitzen, Egbertus Jan, 1881—1966),荷兰数学家.

前面我们提到皮亚诺曲线给线段与正方形之间作成连续对应的事实,嘉当还证明两者之间可以赋予一对一的对应.这么一来我们就要问."什么是维数呢?"特别是,从拓扑学上来看,如何严密证明像线段与正方形之间不可赋予一对一而且连续的对应? 因此就要求拓扑不变的维数(Dimension)的定义.最初给出这种维数定义的,就是布劳威尔的一篇论文.为了证明维数的不变性,布劳威尔运用了所谓连续映射的单纯逼近(Simplexial Approximation)的方法.这是一种对以后拓扑学发展具有巨大影响的方法.在前面提到的 Alexander 关于贝蒂群不变的证明等,也是按照这个方法进行的.这里简要介绍这个方法如下.

设在一个复形 K 到另一个复形 K' 的映射中,K' 的单形 x' 对应于 K 的单形 x,而且 x' 的边界 $\partial x'$ 对应于 x 的边界 ∂x,称这种映射为单形映射.

假设给定了两个曲多面体 P, P' 和从 P 到 P' 的连续映射.我们"细分" P 和 P',做出复形 K 和 K',而且按照适当的单形映射来"逼近"所给定的连续映射.我们可证明这是可能的.这样,从单形映射的组合性质导出连续映射的点集论性质变为可能了.

布劳威尔在这篇论文里处理的点集仍然是曲多面体,但是在第二年关于闭曲线的不变性的论文中,开始做出按照一般的一维有界闭集通过多角形逼近的组合方法处理.这是后来 Alexandroff 奠定完全基础的一般闭集拓扑学的萌芽①.Alexandroff 还运用这套理论到维数论去,收到了很大的成果②.

现在,闭集或者紧统(Compactum,即其上的无限多点必有凝聚点在其上的点集)的拓扑学可以完全按照组合论的方法加以控制.就是说,一般拓扑学的对象可用组合论或代数的方法进行处理.20 世纪 30 年代,Alexander, Kolmogorov 等人还就更一般的列紧(Bicompact)空间导入了一种和普通稍有不同意义的"贝蒂群".

9 抽象代数学方法

在上节中,我们有时引用群论的术语.事实上,无论数学的哪一分科,如果

---

① P. L. Alexandroff: Untersuchungen über Gestalt und Lage abgeschlossener Mengen, Annals of Math. 1928.

② Dimensionstheorie, Math, Annalen, 1932.

没有现代群论的基础概念就不可能深入地对它做出议论来.不但是群论,还有抽象代数学对拓扑学带来很大的影响.这个方法是1920年以后,由女数学家埃米·诺特一派的热心运动而逐渐浸透到数学的各部门中去的,尤其是,对拓扑学有很大的影响.

诺特(Noether, Amalie Emmy, 1882—1935),德国女数学家.

维布伦(Veblen, Oswald, 1880—1960),美国数学家.

莱夫谢茨(Lefschetz, Solomon, 1884—1972),美国数学家.

施赖埃尔(Schreier, Otto, 1901—1929),德国数学家.

冯·诺伊曼(Uon Neumann, John, 1903—1957),匈牙利-美国数学家.

结果之一是"代数复形"的拓展.表示式

$$C = \sum t_i x_i$$

的系数 t_i 不一定限于有理整数,而把它拓展到加法群(Module)的元,也可以做出类似的理论来.原先,维布伦等把它取作为有理整数 mod 2,莱夫谢茨则把它取作为有理数.至于取作为一般加法群的元,那是由于诺特等人的指导思想带来的.其结果是:获得了对于"一般"贝蒂群的有兴趣的一些定理.举一个定理:

有限复形在任意系数域中的各维贝蒂群决定于它在下列各特殊系数域中的各维贝蒂群:

(1)有理整数;(2)有理数;(3)mod m 的剩余系($m=2,3,\cdots$).

证明可参阅 P. Alexandroff-H. Hopf: Topologie I, Kap, V4.(1935).

第二个显著影响来自拓扑代数(Topological Algebra).这个术语是范·丹志希在1932年所取的,它的含义是:把群、环(Ring)、域(Field)等代数学对象同时看成拓扑空间,而研究其构造.范·丹志希最初对"拓扑域"进行了研究,而这方面最有兴趣的对象是拓扑群(Topological Group),它的特殊情况在"拓扑代数"这个名称出现的很早以前就被处理了.古典的李的连续群就是拓扑化的群,Sophus Lie 对这个群解析地表示,假定生成函数的可导微性,而且限于幺元的邻域里进行了研究.施赖埃尔最初用公理法确定连续群的意义,对它作了整体的(Global)处理.此外,希尔伯特在1990年著名的演讲里提出了23个问题,其中第5问题是我们今天所谓的求李群的特征.对此,冯·诺伊曼(1933)得出了决定性的结果:紧的而且局部连通的 n 维拓扑群是李群.Pontrjagin (1934)研究拓扑阿贝尔群的构造,而且得到了重要的结果(参考下节).

10 几个显著的成果

为了补充上述内容的不足,现举出如下的三个例子.

(1)维论(Dimension Theory).我们已提过皮亚诺和嘉当出奇的例子,使人对历来数学中使用的"维"的意义产生了疑问.以后在一段时间里,人们仅仅关于那些能用解析表示的空间乃至图形才谈论维数.对这种状态表示不满的是庞加莱,他根据"分割"(Coupure)递归式地定义了维数.凡是一个被($n-1$)维连续统所分割的连续统,是 n 维,而且对我们种种的感觉空间作了定维数的尝试.读者可参阅著作:*Valeur de Science*(《科学的价值》)第三、四章.



庞加莱的定义有不完备之处.例如,拿在顶点相连接的两个圆锥面(图12.6)来看.它该是2维的,但是按照这里的定义,用过顶点 O 的平面来分割它,得到的是0(0维),于是原图形便是1维了.

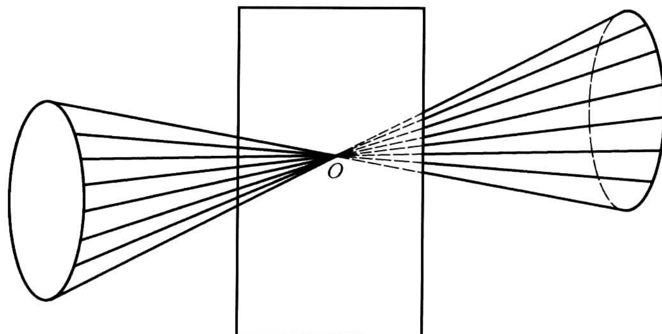


图 12.6

为了弥补定义的缺陷,布劳威尔第一个给出了决定性的定义.现在就连续统 F 叙述布劳威尔的维的定义于下.

设 ε 为任意正数.如果能够用有限个直径不超过 ε 的闭集遮盖 F 的各点而且在不大于 $n+1$ 回遮盖之下能够把 F 遮盖遍的话,那么我们说 F 的维数 $\dim F$ 不超过 n

$$\dim F \leqslant n$$

如果同时不成立

$$\dim F \leqslant n-1$$

我们说

$$\dim F = n$$

另外,如果对于所有的 n 都不成立

$$\dim F \leqslant n$$

我们说

$$\dim F = \infty$$

换句话说,取无论怎样小的"砖头"来对 F 砌墙,使得 $(n+1)$ 个以上的砖头不在一点相会时,这样的最小非负整数 n 就是 $\dim F$.

图12.7示意了正方形(二维)的砌墙,这里大于 $2+1=3$(块)的砖头决不相会于一点.

按这定义得知:同胚的两图形必有同一维数,但是必须证明,已知的 n 维欧氏空间的单形 F 在这个意义下是 n 维的.对此,要证明

$$\dim F \leqslant n \text{ 和 } \dim F > n-1$$

前者是用图 12.7 那样的作图加以证明的. 后者的内容就是:

"如果用充分小砖头砌起 n 维单形,那么无论怎样,一定会产生大于或等于$(n+1)$个的砖头相会的点."这就是所谓"砌墙定理"(Plaster Theorem)①.

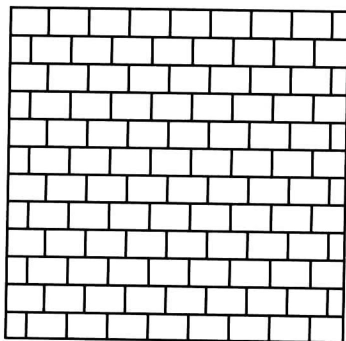


图 12.7

K. Menger 从一般化维数的角度出发,对一般的拓扑空间从而对一般的点集作了如下的定义.

某空间的"维数不超过 n",指的是这样:空间的各点必有维数不超过 $n-1$ 的一个境界,而且被任意小的邻域所包含着. 此外,还假定空集且只有空集才是 -1 维. 这样一来,"维数不超过 n"这句话就递归式地被定义下来了. 至于维数等于 n 的定义则和前述的相同. Menger 立足于这个定义,证明了所谓 Summensatz(和式定理:由维数不超过 n 的可数个闭集作成的和集,它的维数不超过 n)和 Einbettungssatz(嵌入定理:维数不超过 n 的集合与 $2n+1$ 维欧氏空间内的某点集是同胚的),而且完成了他自己的维论. 可是由于他所讨论的对象过于一般化,证明极其复杂困难,所以 Alexandroff 在其维论里只限于讨论欧氏空间内的紧集,再应用前述的论文 *Gestalt und Lage* 中的方法,把 Menger 的结果大大拓展而又严密证明了一些定理. 详细情况可参阅 W. Hurewicz – H. Wallman, Dimension Theory, Princeton, 1941.

(2)对偶定理(Duality Theorem). Alxander 的对偶定理表明了欧氏空间 R 内的曲多面体 P 与其补集 $R-P$ 的贝蒂群之间的关系. 详言之:设 R 是 n 维的,那么一般成立

$$p^{r}(P) = p^{n-r-1}(R-P), \quad 0 < r < n-1$$
$$p^{n-1}(P) = p^{0}(R-P) - 1$$

① "砌墙"的思想是勒贝格所开始的,但严密证明"砌墙定理"的还是布劳威尔,后来 A·施佩纳做了简单证明.



0 维贝蒂数表明了有关曲多面体是由几个连通部分所组成的情况，因此末式意味着下列“分割定理”：

曲多面体 P 将空间 R 分割为 $p^{n-1}(P)+1$ 个连通部分.

这就是若尔当曲线定理最一般的拓展：实际上，当 $n=2$ 时，$p^1(P)=1$，因此欧氏平面被所谓若尔当曲线(圆的拓扑象)分割为二连通部分.

此外，还有庞加莱对偶定理.

设 M 为可定向的流形，它的维数是 n，那么成立

$$p^r(M)=p^{n-r}(M),\quad 0\leqslant r\leqslant n$$

Pontrjagin (1931)注意到这些对偶定理都是建立在同一的代数基础上的事实.同时，他应用 Alexandroff 的理论，把这些定理扩充到一般连续统的场合.此外，还推进了这时所获得的思想方法，发现了拓扑阿贝尔群与其指标群之间成立的“对偶定理”，从而明确了拓扑阿贝尔群的构造①.上述的 Kolmogorov 的“贝蒂群”导入中，也利用了这个结果.Pontrjagin 的对偶定理是属于现代数学的基础定理之一.

(3)连续映射论与不动点定理

关于连续映射的研究中，有所谓布劳威尔的映射度论(参阅前述的 Alexandroff-Hopf：Topologie 第一章 XII).我们在这里不叙述这个理论，只是作为它的应用而导出的不动点定理中，提一提一个直观的特殊情况.例如，下列定理成立：

在单形(任何维)映射到它本身去的连续映射之下，必存在不动点(即映射到它本身的点).

这个“不动点定理”由 Lefschetz 和 Hopf 拓展到一般流形和曲多面体映射到各自本身去的连续映射的场合，而成为决定性的结果(1926 年和 1929 年).另一方面，G.D.伯克霍夫和 Kellogg 把它扩充到函数空间，而且用以表明：它适用于微分方程、积分方程等的解的存在定理的证明(1922)，他们的结果后来为 Tychonoff 等所改进，接着又为 Leray 和 Schauder 所利用，以同样的方法研究椭圆型偏微分方程的解的性质.

读者可参考：江泽涵.拓扑学引论[M].上海：上海科学技术出版社，1979.

① L. Pontrjagin: The Theory of Topological Commutative Groups Annals of Math, 1934(35):361-388.

第十三章 低维拓扑学

1 什么是低维拓扑学

拓扑学是20世纪理论数学中发展最为迅速的分支之一.20世纪60年代是代数拓扑学黄金时代.这样讲不仅是由于这些理论本身的完美,也由于它们已经渗透到数学的许多别的分支中去了.代数拓扑学的工具对于研究维数大于或等于5的流形结构起了本质的作用.许多高难的问题,由于代数拓扑学的发展而变得明朗了.但是对于了解我们所生活的维数的空间.这些工具没有给我们太多的帮助.所以一些数学家们开始问自己:拓扑学是否走得太远了?同时他们也逐渐把自己的兴趣转向2维、3维、4维流形的研究.所谓低维拓扑学,并没有一个很明确的定义.通常的理解就是2维、3维和4维流形的拓扑学研究,它是近年来数学中最活跃的分支之一.

2 早期的低维拓扑学

上述意义的低维拓扑学并不是20世纪60年代才开始的,它可以追溯到19世纪的闭曲面的分类.那时候的低维拓扑学的研究主要关心2维的理论.直到20世纪50年代,数学家们对于3维流形知道得还少得可怜.





有一类 3 维流形，叫做 Seifert 流形，算是了解得比较清楚了．粗略地说．Seifert 流形是这样的一种流形．它们可以看做是一组闭曲线的并，而且这些闭曲线在局部是很规则地排列起来的．例如在一个实心环上，我们可以给出下面的 Seifert 纤维化．一个实心环可以看做是一簇平行的环面加上实心环的中心线组成的．图 13.1 中我们画出了其中一个环面上的两种不同的纤维化．这两种 Seifert 纤维化的不同之处是，图 13.1(b)中的纤维化不是一个乘积纤维化，中心线是一个奇异纤维，因为这个纤维绕实心环的洞是一圈，而它旁边的别的纤维都绕许多圈．如果我们把每一个闭曲线看做一个点，那么这个 Seifert 流形的商则是一个 2 维曲面，换句话说，Seifert 流形是一个用曲面参数化的闭曲线族，用关于这个曲面和那些奇异纤维的信息可以完全刻画 Seifert 流形．

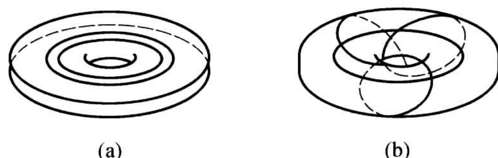


图 13.1

另一个比较早就知道的有趣的定理叫做连通和分解定理．我们先看看类似的 2 维情形，如图 13.2 所示．在一个 2 维曲面 F 上，如果有一条分离的简单闭曲线 C(分离指的是沿这条曲线割开 F，我们将得到两块)，沿这条曲线 C 割开

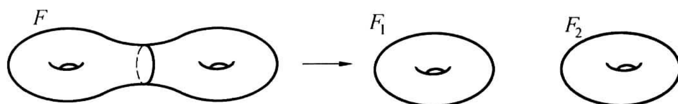


图 13.2

F，我们得到两个带圆周边界的曲面，把这两个圆周边界分别捏成点，我们得到两个曲面 F_1，F_2，这个过程叫做 F 的一个连通和分解．如果我们用 3 维流形中的分离球面，代替 2 维流形中的分离简单闭曲线，我们就得到 3 维流形的连通和分解．如果我们允许得到的流形中可以有新的 3 维球面出现，那么这个分解可以无休止地进行下去．因为我们总可以在 3 维流形中取一个小的实心球，沿其边界球面割开再把边界捏成一点，这小实心球就形成一个新的 3 维球面．有新的 3 维球面出现的连通和分解叫做平凡的分解，连通和分解定理说，对每一个 3 维流形，我们可以做若干次非平凡的连通和分解，使得最后得到的那些流形不再有非平凡的分解，而且本质上这种分解是唯一的．这个定理类似于整数的素因数分解定理，所以我们又称那些不再有非平凡的连通和分解的 3 维流形为素的 3 维流形．根据这个定理．了解 3 维流形，我们只需要了解素的 3 维流形

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

就足够了.

3 20世纪60年代和70年代的组合3维拓扑学

20世纪60年代中期,Haken引入了一个很重要的概念——“足够大的3维流形”.这种流形中,有一张很好的嵌入的不平行于边界的曲面 F(假定它不是球面).而且 F 上的每条可以在这个3维流形中收缩成一点的简单闭曲线.在 F 中也可以收缩成一点,具有这种性质的曲面叫做一张不可压缩的曲面.如果这张曲面是双侧的,那么这张曲面的基本群就嵌入到3维流形的基本群成为子群.这样,这种流形的基本群足够大,包含了一个曲面群,这也是这个名称的由来.通常称足够大的素的3维流形为Haken流形(可以证明:每一个带边界的素的3维流形都是Haken流形),除非它是一个3维实心球.对于Haken流形,我们有一个很好的分类定理:同伦等价的Haken流形是同胚的.

证明这个分类定理,最重要的工具是hierarchy.用2维作对比,我们可以沿一些闭曲线和弧割开一个曲面,使它成为一个多边形.在3维中,我们要沿一些曲面去割3维流形.Haken流形中存在不可压缩的曲面,而且可以证明这些曲面是由它们的基本群在3维流形基本群中的像决定的.沿一张不可压缩曲面割开原来的流形后.我们仍然得到一个Haken流形.或是3维实心球.因为这时候流形带边了,素的条件可以很容易地从这张割开曲面的不可压缩性推出来,我们可以继续这样地割开3维流形.除非我们得到一组3维实心球.关键的一步是证明我们不能无休止地这样做下去,所以有限步以后,我们得到一组3维实心球.这个一步步割开3维流形的过程叫做一个hierarchy.现在我们来看一下分类定理的证明,假定 M,N 是两个同伦等价的3维Haken流形.$f:M\to N$ 是一个同伦等价,固定 N 上的一个hierarchy.第一步是沿一张不可压缩的曲面 S_1 割开 N,得到的流形记做 N_1.注意 f_* 在基本群上是同构,所以 $f_*^{-1}(\pi_1(S_1))\subset\pi_1(M)$ 是一个曲面群,我们可以找到一张不可压缩的曲面 $F_1\subset M$,使得 $\pi_1(F_1)=f_*^{-1}(\pi_1(s_1))$,然后同伦 f 到 f_1,使 $f_1:F_1\to S_1$ 是一个同胚,而且 f_1 诱导出 $M-F_1\to N-S_1$ 上的一个同伦等价,归纳地做下去,我们可以得到一个在两组实心球上的同伦等价,而且在边界上是同胚,所以这个同伦等价可以很容易地用同胚去代替.把这些同胚粘起来,我们就得到一个 M 到 N 上的同胚.

这个阶段中,3维流形的组合研究有了很大的发展.纯粹组合方法的最后一个漂亮的定理叫做环面分解定理,想法有点像连通和分解,不过现在我们是沿一些不可压缩的环面去割一个素的3维流形,割开以后我们得到一些Seifert流形和其他一些不再包含任何不平行于边界的不可压缩环面的流形.有意思的是这种分解还是在本质上唯一的.前面我们已经说过,Seifert流形算是了解得





很好了,余下的是了解那些3维流形,它们不包含任何不平行于边界的不可压缩的环面.

4 瑟斯顿对曲面的研究工作

瑟斯顿(Thurston, William, 1946–2012),美国数学家.

当然,在20世纪70年代以前的3维流形的研究方法也不仅仅是组合方法,例如Novikov的叶结构理论,但主要是组合方法.

研究方法上的真正的发展是20世纪70年代中期开始的,而且最初的那些理论都是和瑟斯顿的名字连在一起的.

瑟斯顿早期的工作是关于流形上的叶结构的存在性的.20世纪70年代中期,他利用一些分析学家们熟悉的概念,例如Teichmuller空间,拟共形映射等,来研究2维曲面上映射的分类,发展了一整套理论,许多搞黎曼曲面的数学家们为之一振,使拟共形映射的研究得到了新的活力.

我们来看看瑟斯顿的主要的方法是什么.首先我们描述一下一个亏格大于1的闭曲面Teichmuller空间.假定F是一个这样的曲面,粗略地讲,F的Teichmuller空间.$\mathscr{T}(F)$是F上的双曲结构的同痕类的全体.它可以用下面的方法构造出来:如图13.3所示,如果F的亏格是g,我们可以找到$3g-3$个简单闭曲线,沿这些曲线割开F,每一块是一个带两个洞的圆盘.在一个带两个洞的圆盘上的双曲结构可以由它的三个测地边界的长度决定.当我们把这些双曲块黏合起来时,我们还需要有一个角度去度量这样的粘法.由于我们考虑的是双曲结构的同痕类,所以图13.4中的两种粘法被认为是不一样的,这样一来每条简单闭曲线提供了两个参数,一个是长度,另一个是黏合的转动角度,或者说沿了$\mathscr{T}(F)$是同胚于R^{6g-6}的.F上的每一个自同胚诱导$\mathscr{T}(F)$上的一个自同胚,而且同痕的自同胚给出$\mathscr{T}(F)$上相同的映射,这个曲面上的自同胚的同痕类构成一个群,通常叫做映射类群.

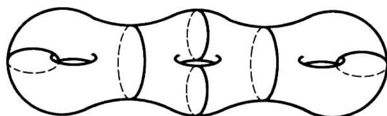


图13.3

Teichmuller空间的结构是很早就清楚的.Thurston的贡献在于他首先给出了$\mathscr{T}(F)$的自然的紧致化,使得映射类群在$\mathscr{T}(F)$上的作用可以延拓到边界上去,他是这样考虑的.$\varphi(F)$表示F上的非平凡的简单闭曲线的同痕类的集合,对于$\mathscr{A}(F)$中的两个元素α,β,定义$i(\alpha,\beta)$是在α,β的同痕类中简单闭曲线的最少相交数.瑟斯顿证明了,对每个$\alpha\in\varphi(F)$,我们可以找到一个

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

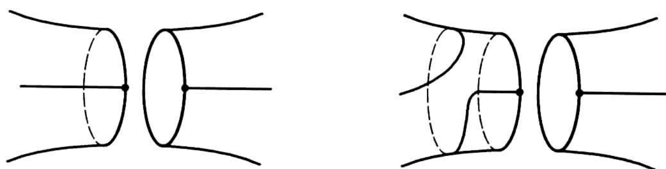


图 13.4

$\beta\in\varphi(F)$,使得 $i(\alpha,\beta)\neq 0$,而且对于两个不同的元素 $\alpha_1,\alpha_2\in\varphi(F)$,我们可以有一个 $\beta\in\varphi(F)$,使得 $i(\alpha_1,\beta)\neq i(\alpha_2,\beta)$.这样一来,$i$ 便定义了一个单的映射 $i_*:\varphi(F)\to\mathscr{P}R^{\varphi(F)}$.而且象的闭包 $P\varphi(F)$是紧致的,这就是 $\mathscr{F}(F)$的一个自然的边界.为了能看清楚这一点,我们可以把 $\mathscr{F}(F)$也嵌入到 $\mathscr{P}R^{\varphi(F)}$中去.对每一个 $\alpha\in\varphi(F)$,$\gamma\in\mathscr{F}(F)$,$i(\gamma,\alpha)$表示 α 中的简单闭曲线在γ 这个几何结构下的长度的下确界,这个 i 有着与前面的 i 几乎相同的性质,同样也给出一个到$\mathscr{P}R^{\varphi(F)}$中的嵌入,象的闭包则是象与 $P\varphi(F)$的并,事实上,我们还可以证明.这个紧致化同胚于一个 $6g-6$ 维的实心球.

这个紧致化的一个直接推论是:F 上的任何一个自同胚,或者在 Teichmuller 空间上有不动点,或者在 $P\varphi(F)$中有不动点.在第一种情况,Teichmuller 空间中的不动点给出 F 上的一个双曲结构,而且这个自同胚同痕于一个等距,第二种情况要分析那些在 $P\varphi(F)$中的不动点.可以证明这时候自同胚根据有一个不动点或两个不动点分成是可约的,或者是伪 Anasov 的.这就是瑟斯顿的关于曲面上自同胚的分类定理.

5　3 维流形上的几何结构

瑟斯顿是一个神秘的人物,他不屑于写他自己做的工作,于是很多人就帮他写,正当人们热衷于写他在曲面上的工作时,他自己却悄悄地开始了对 3 维流形的几何结构的研究.

他证明了如果一个 Haken 流形 M 不包含任何不平行于边界的环面,那么 M 的内部允许一个完备的双曲度量,除非 M 是一个 Seifert 流形或者是一个圆周上的环面丛.证明还是在 hierarchy 上用归纳法.我们来看归纳法中的一个情况,取 M 中的一个不可压缩曲面 F,然后沿 F 割开 M,我们能得到两个允许双曲结构的 3 维流形,想法是把这两个双曲几何粘起来,得到 M 上的一个双曲结构.最简单的情形是我们可以找到两个全测地嵌入的曲面平行于那些黏合的边界,而且黏合映射是一个等距,即我们可以直接把这两个双曲结构粘在一起.一般说来,我们不能保证黏合映射是等距,甚至不是永远可以找到这样的全测地



嵌入曲面的.所以我们要了解怎样去改变这些几何结构,使得它们足够好.可以证明一个双曲流形上的几何结构 1-1 对应于它们在非环面边界上诱导的复结构.所以双曲模空间是一些 Teichmuller 空间的乘积.瑟斯顿用不动点定理证明了,我们总可以形变这两个双曲结构,使得在无穷远的黏合面上的复结构是一致的,而且黏合映射是一个共形映射.这样 Maskit 的定理告诉我们,可以用这两个双曲结构得到 M 上的一个双曲结构.

尽管 Haken 流形算是有了一个分类定理了,不过双曲结构的存在性也还是很重要的.有了这个结果,3 维的拓扑学大大地向前迈了一步.对于 Seifert 流形,我们有 6 种几何学,而且圆周上的环面丛又有一种几何,3 维的情况我们一共有 8 种几何.瑟斯顿根据他的定理,提出了一个猜测:任何 3 维流形,经过连通和分解和环面分解后,得到的块上都有几何结构.这个猜测是 3 维拓扑学中最主要的猜测了,过去的很多著名猜测,包括至今悬而未决的三维的庞加莱猜想,都是瑟斯顿猜测的特别情形.

瑟斯顿还证明了一个关于 3 维流形上的几何结构的结果,如果一个素的 3 维流形允许一个有限群的作用,而且这个作用有大于 1 维的不动点集,那么这个 3 维流形有上面讲的那种几何分解,而且这个群作用限制在分解后得到的每个块上是等距作用.那一段时间是 3 维流形的很光辉的时代.

6 极小曲面的应用

20 世纪 70 年代末和 80 年代初时的另一个热点是极小曲面理论在 3 维拓扑学中的应用.这套方法首先是由丘成桐和 Meeks 开始的.当然,证明极小曲面的存在性也是很困难的,不过那不是低维拓扑学所关心的.他们考虑一个具有最小面积的非平凡的球面,看上去这种曲面应该有最少的自身相交的性质.事实上,他们证明了,如果这个球面同伦于一个嵌入的球面,那么这个最小面积的球面本身也是一个嵌入或者是到一个嵌入的射影平面上的 2 叶覆盖.更有意思的是两个这样的球面或者重合或者相互不交.这样一来,我们可以用这个结果到有限群在 3 维流形上的作用.如果我们有一个有限群 G 作用在 3 维流形 M 上,不难在 M 上给出一个度量,使得 G 是一个等距群.利用极小曲面,我们可以找到一组等变的球面.这样,如果 M 有非平凡的连通和分解,那么我们可以找出一个等变的连通和分解,这叫做等变的球面定理.

极小曲面的理论在等变理论中有许多应用,主要的哲学根据是极小曲面应该有最少相交的性质,也有人研究等变的大亏格的极小曲面.不过在这些研究中,等变的球面定理是最有意义的.

7　单连通闭 4 维流形的分类

前面我们主要讨论 2 维和 3 维流形的研究，4 维流形的早期研究的确很少.假定 M 是一个闭的 4 维流形，那么在第 2 个上同调群的自由部分 $FrH^2(M,Z)$上的上积

$$I_M: FrH^2(M,Z)\times FrH^2(M,Z)\to H^4(M,Z)=Z$$

给出一个整二次型，这个整二次型叫做 M 的 2 次相交形式.

庞加莱对偶定理说这个相交形式是么模的，整二次型当然可以看做是一个实的二次型，所以符号差 $\sigma(I_M)$是一个流形的不变量，通常记作 $\sigma(M)$.整二次型也有一些不同于实二次型的地方，如果对于任何 x，$I_M(x,x)\equiv 0\ (\mathrm{mod}\ 2)$，我们称这个整二次型是偶的.一个二次型是偶的，当且仅当它的矩阵表示的对角线上的元素全是偶数.例如

$$\boldsymbol{H}=\begin{bmatrix}0&1\\1&0\end{bmatrix},\boldsymbol{E}_8=\begin{bmatrix}2&1&&&&&&\\1&2&1&&&&&\\&1&2&1&&&&\\&&1&2&1&&&\\&&&1&2&1&0&1\\&&&&1&2&1&0\\&&&&0&1&2&0\\&&&&1&0&0&2\end{bmatrix}$$

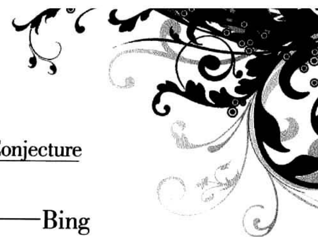
是偶形式.不是偶的二次型叫做奇二次型，奇形式和偶形式在整数环上是永远不会等价的.

人们很早就认识到这个二次相交形式是单连通闭 4 维光滑流形的一个重要不变量，它决定了这种流形的伦型.事实上，它决定了这些流形的 h-协边类.过去人们总认为 4 维是一个最黑暗的维数，距离得到一个好的分类定理太远.1980 年，Freedman 得到一个单连通闭 4 维流形的一个完整的拓扑分类定理.一下子，这个最黑暗的维数变得光明了.他的定理说：从闭 4 维单连通流形 M 到它的相交形式 I_M 的映射是满的，而且是 2 到 1 或 1 到 1 的，依赖于 I_M 是奇形式或偶形式.

惠特尼（Whitney, Hassler, 1907—），美国数学家.

这个定理的证明与高维的类似结果的证明很像.高维的方法不能用到低维上来，主要是因为惠特尼技巧在低维时失效了.1973 年，Casson 有一个方法，可以说是一个同伦范畴中的惠特尼技巧，Casson 当时就意识到，如果这也是一个光滑范畴中的惠特尼技巧的话，那么 4 维可以和高维做得一样好.Freedman 花了很长时间研究 Casson 的这个方法，最后他证明了 Casson 的方法事实上是一





个拓扑范畴的惠特尼技巧,这个证明用了一种比较几何的点集拓扑学——Bing的拓扑学,惠特尼技巧在拓扑的范畴中可以通过,那么 Freedman 的定理就得到了.

8 4维光滑流形拓扑

尽管 Freedman 的定理给出了所有的4维闭的单连通拓扑流形,我们仍然不清楚有多少光滑流形,也不清楚拓扑流形与光滑流形的差别. Rohlin 有一个定理说:单连通的光滑闭 4 维流形 M,如果 I_M 是一个偶形式,那么 $\sigma(M) \equiv 0 \pmod{16}$.初等的整二次型定理说,偶形式的符号差模 8 为 0,这样我们知道在具有偶形式的闭单连通 4 维流形中,至少有一半是不可光滑化的.事实上,在那些具有奇形式的流形中,也至少有一半是不可光滑的,Freedman 的 2 对 1 的映射可以表述成如下的形式,如果我们乘上一个 S^1 得到的一个是可光滑化的 5 维流形,另一个是不可光滑化的,是否还有别的不可光滑的 4 维拓扑流形?

1982 年,Donaldson 在他做研究生的第二年,得到了另一个光滑化的阻碍,他证明了:一个光滑的单连通的闭 4 维流形,如果具有正定的相交形式,那么这个相交形式一定是平方和,我们前面提到的偶形式 E_8 是一个正定的偶形式,但它不能等价于一个平方和,因为平方和是一个奇形式. Rohlin 定理说 E_8 不能是一个单连通闭 4 维流形的相交形式,而 Donaldson 定理说 $E_8 \oplus E_8$ 也不能成为一个单连通闭 4 维流形的相交形式.令人吃惊的不仅仅是这个定理本身和它的很多重要的推论,而且方法上也是很特别的,他把两个原来人们认为不相干的分支——拓扑学和规范场理论——联系在一起了.

固定一个在 M 上的复 2 维向量丛 E,它的第二个 Chern 类是 1.考虑全体自对偶的 Yang-Mills 连络在 $SU(2)$规范变换下的等价类所成的空间 $\mathscr{M}$, Atiyah-Singer 指标定理说,这个 $\mathscr{M}$ 是一个带有限个奇点 $p_1, \cdots, p_s$ 的 5 维光滑流形,每个奇点对应于一个可约的连络.拓扑上,一个可约的连络对应于这样的 2 维同调类 $\pm u \in H^2(M, Z)$, $I_M(u, u) = 1$,所以奇点个数 $s = \frac{1}{2} \# \{u \in H^2(M, Z), I(u, u) = 1\}$.在奇点附近,$\mu$ 是一个 CP^2 上的锥.

我们还需要了解 $\mathscr{M}$ 在无穷远处的情况.这里最基本的工作是 Uhlenbeck 做的.她证明了如下的正则性定理:如果一个序列 $\nabla^i \in \mathscr{M}$ 趋向无穷远处,那么我们找到一个子序列 ∇^{i_j},除了在 M 上的一个点 P 外,∇^{i_j}收敛到一个 M 上的平坦连络,而且 ∇^{i_j}的能量逐渐集中到点 P 的一个越来越小的邻域内. M 上的平坦连络是 1-1 对应于 $\pi_1 M$ 在 $SU(2)$中的表示的,因为 M 是单连通的,我们只有一个平坦的连络,所以能量集中到点 P 的越来越小的邻域中的 Yang-Mills 连络可以用实数来参数化,Taubes 给出了这样的构造.对 M 上的每一点 P,与一个充分小的正数,存在一个这样的能量集中在点 P 的 ε 球中的自对偶 Yang-Mills

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

连络. Donaldson 利用这个构造,证明了在无穷远处,$\mathscr{M}$同胚于 $M \times (0, i)$. 这样我们可以在 $\mathscr{M}$ 中找到一个紧的 5 维流形 $\mathscr{M}$,使得

$$\partial \mathscr{M}_0 = M + pCP^2 + q \cdot (-CP^2)$$

其中 $p + q = s$. 我们知道 $\sigma(\partial \mathscr{M}_0) = 0$,所以

$$\text{rank } H^2(M, Z) = \sigma(M) = q - p \leqslant q + p = s \leqslant \text{rank } H^2(M < Z)$$

这推出 $s = \text{rank } H^2(M, Z)$. 一个初等的代数引理说

$$\frac{1}{2} \# \{u \in H^2(M, Z); I(u, u) = 1\} = \text{rank } H^2(M, Z)$$

当且仅当 $l(\cdot, \cdot)$是一个平方和.

这个定理有一个令人惊奇的推论:$\mathbf{R}^4$ 上存在至少两个不同的光滑结构. 这更显出 4 维是个独特的维数,因为人们早就知道,其他维数的欧氏空间上都只有唯一的光滑结构. 现在人们已经知道 $\mathbf{R}^4$ 上不同的光滑结构多极了,至少要用两个实数来参数化.

Donaldson 是一位大师,他用规范场理论,还给出了 4 维拓扑学中的很多别的结果. 其中有一个是在 $CP^2 \# q(-CP^2)$上有两个不微分同胚的复结构,后来被别的数学家推广为存在无穷多个彼此不微分同胚的复结构. 4 维流形上的光滑结构的理论实在是太丰富了.

9 纽结的 Jones 多项式和 Witten 的工作

低维拓扑学还有一个重要的对象,就是所谓的纽结. 纽结是一个圆周在 3 维球面 S^3 中的拓扑象,而若干个圆周的拓扑象则称为一个链环. 这些拓扑象本身就是一个或一组圆周,没有什么拓扑学可言. 可是研究它们在 S^3 中的放置方法却大有学问. 图 13.5 是最基本的两个不同的纽结,从经验可知它们是不同的,但是如何从数学上来区别它们呢? 最早研究的是它们的补空间,事实上,最近 Gordon 证明了纽结由它们的补空间一意决定. 补空间的基本群就是纽结的一个不变量. 而一般说来,这个群是非交换的. 判定两个非交换群是否同构,是相当困难的,所以这是一个不太好用的不变量. Alexander 曾经得到过一个多项式不变量. 尽管这个多项式没有补空间的基本群所包含的信息多,但是比较两个多项式是否相同太方便了. 20 世纪 80 年代初期 Jones 很偶然地发现了一个纽结的多项式不变量,现在被称为纽结的 Jones 多项式. Jones 是一位分析学家,当



图 13.5



时他的兴趣是有限 Hecke 代数的表示,一次在一个讨论班上,他在黑板上写下一组关系,听众中有人指出,这组关系很像辫群的关系.辫群是一个与纽结关系很密切的东西,此后他潜心研究,终于发现了纽结 Jones 多项式.今天这个多项式已经被拓扑学家弄得十分简单,当时的确是拓扑学上的一件大事,人们没有想到从算子代数的角度可以给出一个纽结的新的不变量.无独有偶,人们一样没有想到,1989 年一个理论物理学家 Witten 用 3 维的规范场理论给出了 Jones 多项式的一个解释.设 K 是 S^3 中的一个纽结,在 S^3 上取一个平凡的 $SU(N)$主丛 E,而且在 E 上放一个连络A.我们可以沿着纽结 K 计算A 的 holonomy,这样得到$SU(N)$中的一个元素,这个元素的共轭类由纽结一意决定,这个元素的迹叫做 Wilson 线

$$W(K) = Tr\, P\exp\int_K A_i \mathrm{d}x^i$$

然后,我们计算 Feynman 道路积分

$$\int D\mathscr{A}\exp(i\mathscr{L})\, W(K)$$

这里 $D\mathscr{A}$表示在全体连络的规范轨道上求的 Feynman 积分,而 Lagrange 作用量取的是 Chern-Simons 形式

$$\mathscr{L} = \frac{k}{4\pi}\int_M Tr(A \wedge \mathrm{d}A + \frac{2}{3}A \wedge A \wedge A)$$

很容易看出,我们这里不用任何度量.自然我们得到的应该是一个拓扑不变量.这个 Feynman 积分依赖于两个变量,一个是作用量中的 k,另一个是群 $SU(N)$中的 N,这两个变量只能取整数值,Witten 证明了,解析开拓 k 和N 到复变量给出两个变量的 Jones 多项式.

Witten 不仅给出 Jones 多项式的这个解释,早前,他有过一个 Morse 不等式的看法,1988 年他还给出一个 4 维超对称规范场模型来解释 Donaldson 的多项式和 Floer 同调.数学与物理在这里这样好地携起手来,实在是很让人振奋的.

低维拓扑学确实是近十几年来数学中最活跃的分支.Thurson 因为他在 2 维、3 维流形上杰出的工作得到了 1982 年的菲尔兹奖,丘成桐的极小曲面的工作也是他得到 1982 年菲尔兹奖的主要工作之一.Freedman 和 Donaldson 因为他们在 4 维流形上的工作得到了 1986 年的菲尔兹奖,而 Jones 和 Witten 则因为他们各自的工作得到了 1990 年菲尔兹奖.特别是理论物理学作为一种工具或作为一种看法,与低维拓扑学还有什么关系,还有很多问题值得去研究.很多大数学家仍在不断地思考.作为现代数学主流中的一部分的低维拓扑学还会继续发展一个时期.

从网络理论到拓扑学①

第十四章

1949年一个美丽的秋日，我（Bott）和我夫人开着我们崭新的Buick牌1934年造的敞篷汽车到一个小山谷去，普林斯顿高等研究所的主楼Fuld大厅就矗立在这个山谷的最前端，这座主楼尽管明显让人联想到Howard Johnson②饭店但却很雄伟，对于我们来说此刻就像是进入天堂一样，我想这部分是由于反差的关系：从匹兹堡到普林斯顿，从三个月前由于疾病和失业而近乎绝望到恢复健康并开始一次奇妙的冒险.

我就从这里开始我的叙述吧，因为这也将很快是我第一次听到拓扑这个词的地方，而拓扑后来成了我主要的专业爱好.我想在这种场合与其像有见解的，即老资格的人们那样去总结这个学科和它的发展，还不如给你们谈一点普林斯顿1949年的事情并讲讲我同这个学科，实际上是整个纯数学的第一次交手.因此这次谈话的真正题目应介于下列题目之间：《从网络理论到拓扑学》或《朝觐Robert Oppenheimer的一位工程师》或《1949~1951年前后在普林斯顿研究所——一个莽撞无知的人的回忆录》.

还是回到1949年的那个下午吧！我可以自豪地说我们欣喜的心情甚至在见到我们在“大院”里的公寓时也没有减退，这里的大院是指供“临时成员”居住的废弃的兵营.当然喜欢寒冬半夜熄火的煤炉其缺点在9月份还显不出来.

① 原题：On Topology and Other Things.原作者：Raoul Bott.本文是以1984年8月在Eugene举行的暑期会议中的一次谈话为基础，这次会议是美国数学学会和美国数学协会联合主办的系列历史讲座之一.——译注

② 美国一家大众化的同名旅店和餐馆的创始人，分店遍设全美国各地.





至少对我们来说那整个第一年确实是令人陶醉的,我们得不时地掐掐自己以确信这一切都是真的,如果换作你,你也会这样做的.想想这样一个地方吧:警察想要逮捕的形迹可疑的流浪汉结果是勒雷;在这儿每天早上11点左右,要与爱因斯坦就诸如天气或邮件递送的迟缓这类似乎重要的话题闲谈是很容易的;在这儿,在一群喧闹的吃午餐的年轻人中,一个友好但一句话也不说的邻座原来是迪拉克等等.我在这儿所描述的是年轻人身处他们的偶像之中的那种兴奋,而且当他们了解到偶像们的怪癖后那种兴奋和鼓舞只会是增加.我相信这种现象今天仍在继续,而且在我们之中拥有像普林斯顿这样的研究所的巨大优点之一——就是在那儿所引起的那种兴奋,那种兴奋使得年轻的和年老的团结在一起,从而使年轻人牢牢地与他们的学科连在一起.然而我还是禁不住被权威的"老人"的幻想所迷住.他认为在他那年代里情形总要美妙些;因此我不禁觉得那些过去的日子才是普林斯顿研究所真正庄严的日子维布伦、亚历山大、莫尔斯、外尔、冯·诺依曼,以及西格尔是数学正教授;塞尔伯格、蒙哥马利和哥德尔是终生成员,而莱夫谢茨、阿廷和博赫纳仍活跃于大学里.Oppenheimer新近被任命为所长,他的作用随处可见,这特别是在饮茶后,通常他简直像赶牲口似地把年轻的物理研究人员们赶到他们的讨论班去.

正如今天数学家们是一个不怎么均匀的群体一样,物理学家们也有各自的工作和玩乐的节律,这些物理学家和我那群同龄人坚信酒精对疲倦的数学头脑的治疗作用.在蒙哥马利家里进行的一段成功的疗程里还有这样一段轶事,鸡尾酒会的某些情节我现在已记不清了,这无疑是由于已过去了很长一段时间的原因.但我清楚地记得最后我们中的几个在蒙哥马利家的地毯上玩打弹子游戏,冯·诺依曼参加了那次酒会.在打弹子的过程中不知怎的我向他问做一个"伟大的数学家"是个什么感觉.以他特有的深思熟虑的风格,他应该认真对待我的问题——尽管刚才他才从他那似乎是各色各样的故事集里一样抽了一点给我们听听.他所说的基本上是这样:完全老实地说,他只知道唯一的一个"伟大的数学家"——希尔伯特,而就他而言,由于是神童,他从未真正感到他达到了人们对他的期望;因此你可以明白被认为不够格并不难,只是需要一个适当的标准而已.

不过现在是说说这次谈话的"严肃话题"的时候了:"拓扑学"或"位置分析"(Analysis Situs).如维布伦、莫尔斯、西格尔和外尔这些人当时所称呼的那样,从前的位置分析现在以"代数拓扑"这样一个笨拙的标题来称呼,它的目标是构造和探讨或多或少可计算的、用来区分不同几何形状的——更一般地说是区分不同拓扑空间的不变量.

从直观上来说,一个空间所分成的块数当然是最明显的不变量,而且在某种意义上说,这个学科整个巨大的大厦依赖于对分支概念如何理解而分成两个分

勒雷(Leray, Jean, 1906—),法国数学家.

迪拉克(Diral, Paul Adrien Maurice, 1902—1984),英国物理学家.

维布伦(Veblen, Oswald, 1880—1960),美国数学家.

亚历山大(Alexander, James Waddell, 1888—1971),美国数学家.

博赫纳(Bochner, Salomen, 1899—1982),德国-美国数学家.

塞尔伯格(Selberg, Atle, 1917—),挪威-美国数学家.

蒙哥马利(Montgomery, Deane, 1909—),美国数学家.

莱夫谢茨(Lefschetz, Solomen, 1884—1972),美国数学家.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

科.如果我们用弧分支的概念,即一个拓扑空间 X 的两点 p 和 q 在同一个分支里当且仅当存在一个映射(图 14.1),即连续函数 $\mu: I\to X$ 从区间 I 到空间 X,使得

$$\mu(0)=p,\mu(1)=q$$

那么这个概念自然地将我们引入"同伦论",于是在同伦论中我们有:

另一方面,如果我们采取这样的主张:两点 p 和 q 在 X 的同一分支中当且仅当所有局部常值函数 f,比如说实值的[1],在 p 和 q 处相同,那么我们将按照上同调论的传统来开始:即如果我们让 $H^0(X)$ 等于局部常值函数的空间,那么在上同调论中:分支个数等于 $\dim_R H^0(X,R)$.

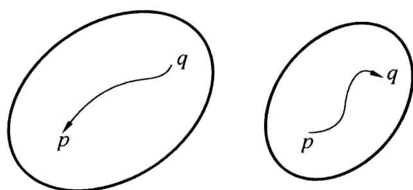


图 14.1　$\pi_0(X)=X$ 中弧分支的集合

在这两种理论中,高阶不变量经某种归纳步骤由其零阶不变量产生,对同伦论来说这步骤是一个相当几何而朴实的步骤;而对上同调论来说则是更具代数性且微妙得多的步骤;在另一方面,实际情况是分支的上同调扩张较容易计算,而同伦群 $\pi_k(X)$ 虽然在概念上较容易但却难以计算.事实上,它们可如下来归纳定义设 p 是 X 的一点,用 ΩX(图 14.2)来表示由所有这样映射 $\mu: I\to X$ 且 $\mu(0)=\mu(1)=p$ 构成的空间,比如说它具有紧开拓扑,我们称 ΩX 为 X 上基于点 p 的圈(Loop)空间.然后如下归纳定义 $\pi_k(X)$

$$\pi_k(X)=\pi_{(k-1)}(\Omega X)=\pi_{(k-2)}(\Omega\Omega X)\cdots=\pi_0(\Omega\cdots\Omega X)$$

这个定义可以追溯到胡雷维奇的定义,而且容易看出它是等价于 20 世纪 30 年代 Czech 的定义的.

胡雷维奇(Hurewicz, Witold, 1904—1957),波兰数学家.

为了使你们了解一下 1949 年时确定这些群所具有的水平,让我告诉你们这样一件事.一天斯廷罗德走进他在普林斯顿主持的纤维丛班——这个班我是作为"第一个拓扑课程"去参加的——并骄傲地宣布:Pontrjagin 刚来信,大意是说 $\pi_4(S^2)$ 等于 0(图 14.3).几个星期后从剑桥传来了兴奋的消息:George Whitehead 却得 $\pi_4(S^2)$ 等于 $Z/2Z$,而且很快 Pontrjagin 证实了 George 的结果.

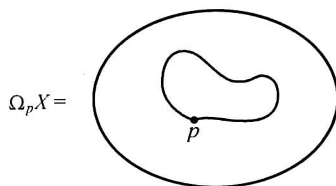


图 14.2

斯廷罗德(Steenrod, Norman Earl, 1910—),美国数学家.

但是 1949 年在普林斯顿并不真正是一个同伦年,倒是在这一年里,在外尔的讨论班上分别由 DeRham 和 Kodaira 给出了 Hodge 理论的两个不同但大概真

[1] 原文为:"……常值函数 f,比如说实值的,和 X 在 p 和 q 处相同."——译注



是无懈可击的证明，除了斯廷罗德在大学的课程外，这个讨论班是最引起我兴趣的一个，这当然是由于网络理论的基本存在定理可看做是 0 维 Hodge 理论的有限维类推！要讲清这一点，首先让我回忆一下分支的上同调扩张概念是如何起作用的——至少是在最吸引我的范畴“光滑空间的范畴”里——即 C^∞ 流形 M 里. 通常由定义，这样一个流形是若干个 $\mathbf{R}^n$ 光滑地粘在一起，因此分析的概念和构造自然地由 $\mathbf{R}^n$ 扩张到这些个 $\mathbf{R}^n$ 上. 特别地，人们可以从这样的 C^∞ 函数集 $A^0(M)$ 着手，其中的局部常值函数相对于任何局部坐标 x 都由下列微分方程所刻画

图 14.3

$$\mathrm{d}f=0 \Leftrightarrow \frac{\partial f}{\partial x^2}=0$$

在这种背景下，我们可以发现 M 的分支个数等于微分方程 $\mathrm{d}f=0$ 的解空间的维数.

而这个 H^∞ 的上同调归纳推广是 DeRham 复形

$$A^0(M)\xrightarrow{d}A^1(M)\xrightarrow{d}A^2(M)\xrightarrow{d}\cdots$$

其中 $A^1(M)$ 表示 M 上的“共变”交错张量场，而“d”是已碰见过的将函数 f 变成它的微分 $\mathrm{d}f$ 的“d”的唯一自然扩张，那么 M 的 DeRham 上同调可简单地由下面给出

$$H^q(M)=\mathrm{Ker}\ d/\mathrm{Image}\ d(\text{在维数为 } q \text{ 的情形})$$

如果 M 还给定了一个正定度量，那么每一个向量空间 $A^i(M)$ 都继承一个自然内积 (ϕ,χ)，而且相应地每一个微分算子“d”有一个确定的共轭“d *”满足通常的共轭方程：$(\mathrm{d}\phi,\chi)=(\phi,\mathrm{d}^*\chi)$. 这样 Hodge 理论就要研究算子 $\square=\mathrm{dd}^*+\mathrm{d}^*\mathrm{d}$ 和它的解空间 $H^*(M)$——M 上的调和形式.

算子 $\square$ 显然是保持 $A^i(M)$ 不变的，而 Hodge 定理断言调和形式恰好给出了任何紧定向流形 M 上的上同调：即，我们有等式 $H^q(M)\simeq H_q(M)$.

现在将这种情形同网络理论的最初的也是最简单的问题进行一下比较，在这里我们研究的是电阻（Risistor）R 的集合，这些电阻是按照一个图（Graph）把它们钩起来——如图 14.4 中所示.

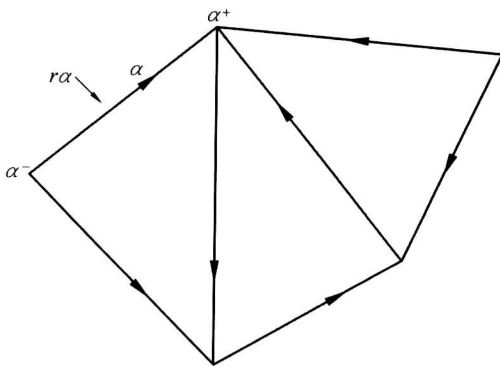


图 14.4

如果$\{\alpha\}$跑遍 N 的分支——这些分支我们以某种固定但任意的方式为其定向——而且如果我们用 α^+ 和 α^- 来代表分支 α 的两个端点,那么“关联”数据 $\alpha \to \begin{pmatrix} \alpha^+ \\ \alpha^- \end{pmatrix}$ 完全决定了 N 的拓扑.用线性代数的术语来说这个关联关系相应由下列线性算子唯一确定

$$\delta : A^0(N) - A^1(N)$$

此算子使定义在 N 的端点上的实值函数空间 $A^0(N)$ 映到定义在定向分支上的函数空间 $A^1(N)$,它由下面式子给出

$$\delta\phi(\alpha) = \phi(\alpha^+) - \phi(\alpha^-)$$

显然这个 δ 正是前面的 d 的有限差分类推,而且正如前面一样 N 的“组合上同调”也可定义如下

$$H^0(N) = \mathrm{Ker}\ \delta, H^1(N) = A^1(N)/\delta \text{ 的象}$$

这些向量空间证明是刚才仅考虑为拓扑空间的那个图(Graph)的拓扑不变量,而且的确很显然的是,δ 的核由定义在 N 的端点上的那些函数组成,这些函数在每一个连通分支是常量,因此很显然我们有关系式

$$\dim H^0(N) = N \text{ 的连通分支的个数}$$

一旦分支的正阻抗(The Positive Resistance)r_a 给定,N 的电学性能当然是唯一确定的.数学上的正确程序是要将倒数 $g_a = 1/r_a$ 解释为由下列公式确定 $A^1(N)$ 上的一个正定内积

$$(\phi\psi) = \sum g_a\phi(a) \cdot \psi(a)$$

并同时赋予 $A^0(N)$“克罗内克”内积

$$(f,g) = \sum_p f(p_1)g(p_1)$$

正如前面一样谈论 δ 的共轭——记为 δ^*——以及相应的 Hodge 拉普拉斯算子$\square = \delta^*\delta + \delta\delta^*$ 是有意义的,而且根据这些,N 的稳态位势分布(Steady-state Potential Distribution)E 由以下的方程而与某个外电流(Exterior Current Flow) I 有关:

$\square_0 E = I$,这里$\square_0 = \delta^*\delta$ 是 0 维的 Hodge 算子.

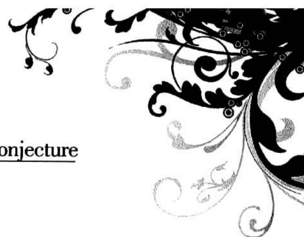
因此在这方面由“Hodge 理论”有

$$\dim \mathrm{Ker}\square_0 = \dim H^0(N) = N \text{ 的分支个数}$$

因而马上证明了网络理论的第一基本存在定理——这个定理对于任何称职的工程师来说当然是自明的——即,$\square_0 E = I$ 对 E 有唯一的解,只要 N 的每个分支在一个端点是“有根据的(Grounded)且流进 N 的任何分支的总电流(Total Current)是零”.

顺便说,这个有限维的 Hodge 理论直接就是线性代数.在另一方面,下面的





内容,更确切地说,在 1949 年开始解出这些方程确实是我引以为自豪和快乐的事,而且我想,这实质上也是我证明的第一个定理.

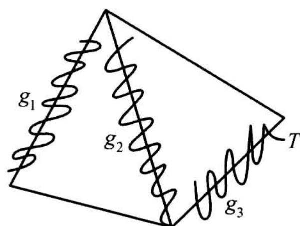


图 14.5

这个定理就是下面这个关于算子□的非零特征值乘积的公式

$$\det\Box'_0 = \prod_{\lambda>0}\lambda$$

对于 N 的每一个(生成树 =)极大树,如图 14.5 所示,设 Tg 等于对应于 T 的分支的那些 g 的乘积 $Tg = g_{\alpha_1}\cdots g_{\alpha_k}$.有了这些我们有 $\det\Box'_0 = \sum^{T} Tg$,这里 T 跑遍 N 的极大树.

我把这行列式 $\det\Box'_0$ 叫做网络 N 的阻抗位势 ψ,由这个行列式,$\Box_0$ 可以在它的零空间的补空间上求其显示表达的逆.特别地

$$\frac{\partial\log\psi}{\partial g_a} = R_a$$

用 ψ 给出了 N 穿过分支 α 的有效阻抗.

唉!——后来我才知道——在这些公式上我实际上被 Kirckhoff 抢先了.我提起这些公式只是由于这个行列式的无限维类推,即狭义 Hodge 理论中的 $\det'\Box_{k'}$ 至今在现代物理学中还起着中心作用,而我在准备这次报告时不禁为这个阻抗后势是否在无限维情形有某种推广而感到困惑.

我也不知道有关 $\det(\Box_0 + \lambda I)$ 的所有系数的公式会是什么样,而我希望它们只包含正树的积.

当然在无限维情形下,$A^k(M)$ 上的每个拉普拉斯算子 $\Box_{(k)}$ 是一个二阶微分算子,因此一点也不清楚它的行列式应如何定义.然而在紧流形上这样的□拥有离散的特征值集,而且在 1949 年时有关特征值的渐近分布已了解得足够多了(特征值的渐近分布是外尔另一个最喜爱的学科)可以对□定义某种正规的行列式概念.

我间接提及的定理是我在外尔关于“新老特征值问题”的演讲中首次听到的,它属于 Minakshisundaram 和 Pleijel,它断言:ζ 函数 $\zeta_s[\Box] = \sum\lambda^{-s}$ (λ 跑遍□的特征值)不仅对实部充分大时有意义而且到整个复 s – 平面上有一个亚纯扩张,而它的极点的位置可以显式地描述出来,特别地 $\zeta_s(\Box)$ 在 $S=0$ 附近是解析的,因此形式上的计算

$$\frac{\mathrm{d}}{\mathrm{d}S}\zeta_s(\Box)0 = |\sum - \log\lambda\cdot\lambda^s|_0 = -\sum\log\lambda = -\log\det\Box$$

可以提供一个正规化的行列式 $\det\Box$:$\det\Box = \mathrm{e}^{-\mathrm{d}/\mathrm{d}s\zeta(s)}|_{s=0}$,而这就是这些行列

赖德迈斯特(Reidemeistor, Kurf werner Friedrich, 1893—1971),德国数学家.

式进入数学文献的经过——这在 Ray Singer20 世纪在 50 年代关于"赖德迈斯特挠率"的文章中特别明显,它们也逐渐地以这种形式出现在场论的当代文献中.

赖德迈斯特这个名字使我回想起 1949 年的那些日子,因为赖德迈斯特和 E. Specker 那年是我在研究所的导师,赖德迈斯特是一个热情迷人而文雅的,具有哲学家气质的数学家,他不管是谈话还是讲课都很流利地使用他自己的语言——大约 40%的英语和 60%的德语——而他在嘉当的"蟹壳理论"(Theory of Cavapace)小型讨论班的讲课——这课程嘉当曾于 1949 年在 Harvard 讲过——一直是我生活中最使人鼓舞但也是令我不解的经历.

赖德迈斯特还给我们这帮年轻的活宝中的一些人传授一种古老的——现在恐怕已消亡的——研究所传统.这些年轻人组成伐木队,星期三下午 2 点准时启程,足迹遍及研究所的占地,伐倒老树或者将倒下的树分块锯得容易处理.参加者有像西格尔,Pauli,Alexander,Alexandraff,赖德迈斯特等.所有这些人在我看来都是了不起的,德高望重的智者——总之,是老资格的——但细想一下那时他们实际上 50 才刚出头,比我现在可年轻多了,唉!

尽管我快把时间用完了,我还是忍不住要讲有关这些伐木活动的(许多件中的)一件轶事.11 月份的一个星期三天正下着雨夹雪,我很自然地断定伐木是不可能的了.于是我想知道西格尔会怎么做.我为这种好奇所驱使,在下午 2 点前几分钟到了我们碰面的地方.果然下午 2 点整西格尔来了.然后我想既然只有我们两人来了,他可能会要求停一次.根本没那回事.我们拿起一个大锯,把它搬到树林里并开始干活——大部分时间里不讲一句话,全都淋得湿透了.在短休时,西格尔突然转过身来对我说:"你也知道,我真羡慕那些今天没来的人.但是一个人怎么能定了在星期三下午 2 点来伐木却仅仅因为天气糟糕而待在家里!"

如此讲原则的日子到哪里去啦!

Carl Ludwig 在我看来是研究所里王子中的王子,他讲课很生动,听他的课你会感觉接触到了数学的每一个分支,从数论开始,你可以碰到分析、拓扑、代数几何还有 p 进数,这一切全在一个小时内!尽管对纯数学的初学者这些全都难以理解,但它们确实比其他任何东西更激励我希望去了解这整座美丽的大厦,特别是了解拓扑在这个学科中所起的作用.我今天仍然认为拓扑不是孤立的而是与分析、几何实际上与整个数学相互作用的,我想我这种见解应归因于 Carl Ludwig 的影响.

当然那时普林斯顿的许多伟人们也持有这种见解.外尔、莫尔斯、博赫纳和莱夫谢茨无疑都表示过那方面的见解,实际上他们对 Eilenberg-Steenrod-MacLane 学派提出"新代数拓扑"是有些敌意的.在这种相互矛盾的压力下,我只好这样做来求得折中:一方面去听 Steenrod 的课——同时在他身旁花上尽可能多的时





间——另一方面是学“莫尔斯理论”——很大程度上靠自学，因为莫尔斯那时对他早期的理论几乎没什么兴趣，但他正同 Transue 专心于有界变分函数这类问题的工作.

让我给你们回忆一下：在莫尔斯理论中，维数 b_k = 一个紧流形的上同调群的维数 dim $H^k(X)$与光滑函数 f 的极值或 M 有关. 事实上，一般情形下这些极值都出现在孤立点，我们让 $m_k(f)$表示恰有 k 个“速降”方向的 f 的极值. 那么 f 的莫尔斯级数

$$Mt(f) = \sum m_k t^k$$

根据一组不等式与 M 的庞加莱级数

$$P_t(M) = \sum \dim H^k(M) t^k$$

有关，关于这一点莫尔斯理论明白地断言

$$M_t(f) - P_t(M) = (1+t)Q(t)$$

这里 $Q(t)$是一个具有非负系数的多项式.

从这些不等式马上有：如果对于奇数 k，m_k 是 0，那么

$$\dim H^k(x) = m_k(f)$$

即知道 f 的极值二阶性质就可推出整个加法上同调结构.

莫尔斯当然把这一原则应用到一个更大的背景，即应用到函数空间，特别是圈空间 Q, M，因此在 1927 年他已对 N 维球 S^n 算出了这圈空间的上同调，他的结果是

$$H^k(\Omega_p S^k) = \begin{cases} Z, \text{当 } k = n-1 \\ 0, \text{其他} \end{cases}$$

而且由此他推出一个漂亮的结果：在任何黎曼结构中一定有无限多条测地线连接 S^n 上任意两点 p 和 q.

从一开始，我就对 $H^k(\Omega S^n)$的计算非常感兴趣，但我确实不知道如何弄清它的来龙去脉，我想当时在普林斯顿，我们中没有人知道，尽管我期望在巴黎的 Serre 那时已经知道了. 整个来说当时真正的拓扑学家不喜欢莫尔斯理论就像莫尔斯不喜欢他们一样. 1949 年时我们正好是处在暴风雨前的宁静之中. 代数拓扑，特别是同伦论里的法国革命还没有冲击到我们. 因此当时不知道有什么严格限于拓扑的方法来得到莫尔斯公式，从而也没有这样的方法来使用我作为一种计算手段介绍给你们的 π_k 的归纳定义，因此在 1949 年关于球的高阶同伦群. 人们确实有把握的只是：它们是可数的.

然而，回顾一下，正像我们今天所知道的那样，除了两个方面外，实际上在当时舞台已为这整个学科搭好了. 我们从 Eilenberg 和 Mac Lane 那里知道看问题的正确方法是应用函数的方法. 我们的先辈所谓的拓扑不变量总是与空间范

畴中的函数以及这些空间到代数范畴(比如群的范畴)的映射有关.这也就是说,在任何拓扑不变量后面都潜藏着一种方法,这种方法不仅赋予每个空间 X 一个群 $F(X)$,而且每个映射 $h:X\rightarrow Y$,有一从 $F(X)$ 到 $F(Y)$(在共变情形)的同态 $F(h)$,比如同伦,在反变情形正相反,比如上同调.我也知道对于几何中的许多经典问题——比如流形上没有零点的向量场的存在性——两个空间的挠积概念是不可缺少的,而且挠曲的量度牵涉同伦论.

这个领域的想法,就其极强的几何意义背景而言,很大程度上分别来自 H. Hopf和 H. Whitriey.我建议每个学拓扑的年轻人有时应该看看 20 世纪 40 年代早期的密西根拓扑讨论会报告(Michigan Colloquium in Topology).那里面的论文看来都是里程碑.特别是惠特尼的文章.在这些文章里,我们现在所称的纤维丛理论已给出其概述.

我们还从 Eilenberg Mac Lane 以及独立地从 Hopf 那里知道:如果空间 Y 只有一个非平凡的同伦群,比如说 $\pi_k(Y)=\pi$(且所有 $\pi(Y)=0, r\neq k$),那么不计与所考虑的空间范畴有关的某种形式的同伦等价,空间 Y 由 π 所决定.因此对每个可换群 π 和整数 n 都存在一个确定的空间,Eilenberg-MacLane 空间 $K(\pi, n)$这个空间由下列的方程来刻画其特征

$$\pi_k\{K(\pi,n)\}=\pi,\text{所有其他的 }\pi_k=0$$

在另一方面,人们也知道:一般说空间的同伦型不能由它的同伦群单独决定,即不是每个空间都是对应于其每个同伦群的 Eilenlerg-MacLane 空间的乘积.空间的同伦型是这些 Eilenberg-MacLane 空间的某种挠积

$$X=K(\pi_1,1)\times_T K(\pi_2,2)\times_T\cdots$$

且每一级(Stage)的挠曲是 H^q((前一级);π_q)的一个元素来测度的.

但所有这些深刻而富有创造性的见解离实际应用相去很远,因为我们确实没有什么好方法来计算挠积 $X\times_T Y$ 的上同调,也没有一个确实好的挠积的定义.

从一方面来说,后来这两个鸿沟被 Leray 谱序列和 Serre 对 Leray 的结果的见解填平,在这儿 Leray 谱序列对挠积的上同调与笛卡儿积的上同调的差异程度给出了确定的界线,而 Serre 的见解是来自于微分几何上的考虑,可以用到非常一般形式的挠积——Serre 的纤维空间.特别地在其领域里他很清楚地看到:圈空间 ΩX 和 X 能够扭在一起(Twisted Together)来产生一个可缩成一点的空间

$$\text{点}=X\times_T\Omega X$$

另外作为诸如球这种简单情形的一个结果,用于这个方程的 Leray 法则足以利用 X 计算出 ΩX 的上同调.因此 Serre 重新证明了莫尔斯公式,但是是用一种他可以反复使用的方法,从而他带来一个时代,在这个时代里人们可以希望只需用 $H^0(\Omega^q X)$.这样的上同调手段就可计算出 $\pi_q(X)$.尽管在任何个别的情形这





仍然是难以办到的,但这种观点从根本上改变了问题.

今天我们知道,例如 $\pi_q(X)$在技术意义上是可计算的(E. Brown 的一个定理).如果 X 是一个有限单纯复形,利用这种方法塞尔当然立刻导出了球 S^n 的高阶同伦群①的有限性.

塞尔(Serre, Team-Pierre, 1926—),法国数学家.

然而我想如果让我们这些听斯廷罗德讲课的学生去注视水晶球来占卜看看会发生什么事,那么最使我们感到惊奇的是拓扑里即将出现的真正几何方面的惊人发展.当然在普林斯顿有一个很活跃的低维"纯正"拓扑学派:Papakiriakopolous,莫尔斯,Fox,但谁会想 5 年后托姆的配边,10 年后米尔诺的怪球及斯梅尔的浸入和环柄体理论,也不会想到 Hauptvermu(主猜测)在高维情形的框架,就更不用说 Kirby 和 Friedman 近来的结果了.从根本上来说,尽管同伦论的基本问题已取得了进展,但是对能否证明纯正拓扑的分类问题仍是悲观的.这种悲观思想是如此之普遍,竟让 J·H·C·怀特海这样出色的几何学家不去干他最擅长的事,而是把时间浪费在这个学科中各种代数较强的方面.

托姆(Thom, Reré, 1923—),法国数学家.

怀特海,"耶稣使怀特海糊涂"我们通常亲切地这样说他.他身上有着一种多么奇妙的魅力! 当他到普林斯顿作短暂访问——是 1949 年还是 1950 年,我不清楚——整个普林斯顿的脉搏立刻快了一倍.我清楚地记得在斯廷罗德家的一次聚会,当时 Henry 向我们大家叙述他如何为 Norman 1945 年那篇见解独到的文章而激动,在这篇文章里斯廷罗德引入了他的 v_i 积——后来称为斯廷罗德平方——而且立刻应用到一个障碍性问题的构造上.斯廷罗德说他只是随便考虑上闭链公式时才偶然地发现它的,而怀特海需要细节.怀特海在夸奖别人时是非常慷慨大方的,但反驳别人时也是十分坦率的.他是冯·诺依曼的老朋友,但对冯·诺依曼卷入政府事务持悲观看法."一个人怎么能为了乏味的权力阶层而抛弃了精神世界呢!"他有一次这样劝导我.

斯廷罗德平方在那时当然还处在发展阶段——而我的每一篇拓扑论文就是研究这些平方以及它们与对称积和 Smith 理论的关系.正像对示性类理论一样,我们大家在研究斯廷罗德平方时都小心谨慎且充满敬意的.毕竟,在惠特尼关于示性类的会议录通告中,他说我们所谓的惠特尼对偶定理是他所证明的最难的定理.

希策布鲁赫(Hirzebruch, Friedrich Ernst Peter, 1927—),德国数学家.

因此我不禁想到事情会是多奇妙,要是与 Henry 在 Norman 家厨房里闲聊时我能给他们讲讲此领域后来出现的一些事情让他们高兴高兴,诸如,"将此定理(惠特尼对偶定理)作为此理论的起始公理要有利得多"或"一个叫希策布鲁赫的年轻家伙在 6 年后竟厚颜无耻地置它们(斯廷罗德平方)于不顾".

① 原文为同调群.——译者注

当然我必须承认如果那天晚上我用经六七年考虑才偶然发现的一个应用莫尔斯理论的结果来让 Henry 大吃一惊的话,那会更有趣.然而那只是可能的玩笑而已,并非真正有那么一回事.不过简短地回忆一下这个(莫尔斯理论的)应用以及我是如何偶然想到的,这样来结束这次讲座也许是适宜的.

当我于 1955~1956 年回到普林斯顿高级研究所时情况发生了很大的变化——塞尔、小平邦彦、Spencer、阿蒂亚、波莱尔、米尔诺、西格尔和希策布鲁赫都成了我的老师和朋友,而同伦论在我的心目中已不是在顶峰位置.我已经注意到莫尔斯理论在李群的拓扑学中有很漂亮的应用,而且人们可用一种很简单的方式,根据 G 的基本数据写出圈空间 ΩG 的同调.但是正如我说过的,1956 年我正在跟塞尔和 Compan 学黎曼定理这一类东西.同时在同伦论研究者与像波莱尔和希策布鲁赫这样的示性类奇才之间产生了争论.1949 年时人们已知道典型群:$U(n)$,$O(n)$,$Sp(n)$的第 k 个同伦群在 n 充分大时不依赖于 k,这些稳定群 $\pi_r(U)$,$\pi_k(O)$,$\pi_k(Sp)$开始成为争论的内容.

小平邦彦(Kodaira, kunikiko, 1915—),日本数学家.

波莱尔(Borel, Armand, 1923—),法国数学家.

阿蒂亚(Atiyah, Michael Francis, 1929—),英国数学家.

同伦论的研究者们预言 $\pi_{10}(U)=Z_3$,波莱尔和希策布鲁赫说不是,说这是不可能的.一定有 $\pi_{10}(U)=0$.碰巧的是,Samelson 和我发展起来的莫尔斯理论观点使我能做一个很独立的计算——实际上是在例外群 G_2 里——来得到了 $\pi_{10}(U)$.计算花了一个星期,在这一星期里我同我亲爱的朋友 Arnold Shapiro 一起.为了计算 G_2 的圈空间的斯廷罗德平方数及其之类的东西,我们把研究所里所有的黑板全占用了.最后终于算完了,结果是我们站在波莱尔和希策布鲁赫一边,而且我还记得我们以 5 美元赌他们是对的,并及时地从 John Moore 那里收到了钱.

一经排除了 $\pi_{10}(U)$等于 Z_3 的可能,人们已知道同伦 $\pi_k(U)$对于 k 为奇数时其形式为 Z,对于 k 为小于 10 的偶数时为 0.

对于像我这样一个老工程师来说这就足够了!我断定,从头到尾的模式一定是这样的.但是当时标准的计算方法看来是没有希望的.在另一方面如果人们能证明 $\Omega \cdot \Omega U=U$,那么由归纳法所有的 k 立刻可得到以上的结果.就此而论,从我以前的工作再加上 Samelson 利用莫尔斯理论结果要推出这个定理并不困难,实际上我只是在 1956 年秋回到 Ann Arbor 时才领悟到这个定理,当时我们正在搬家,我发觉同样的方法似乎可以给出下列方程

$$\Omega^{(4)}0=SP \quad 且 \quad \Omega^{(4)}SP=0$$

从而典型群的所有稳定同伦群实际上都是周期的,特别地

$$\pi_{k+2}(U)=\pi_k(U)$$

$$\pi_{k+4}(O)=\pi_k(SP)$$

$$\pi_{k+4}(SP)=\pi_k(O)$$

但是现在我的时间已真正用完了!Graham Greene 常把它的作品分成“小



说"和"消遣部分".我宁愿把这段个人经历放在后面一类.如果我的同辈人中有足够多的人干同样的事,那么我们确实可以共同重建美国数学独特而又罗曼蒂克的时代,非常感谢你们.

基本群和同调群的直观描述①

第十五章

1 引 言

笛卡儿创立解析几何在数学发展史上是一个伟大的贡献.这个贡献不在于数学上发现了某个重要结果,而在于它是方法论上的一个伟大创造,把数学上的两个大的方面即数与形结合起来,开创了通过坐标系将几何中的图形用代数中的方程表示出来,从而用代数工具研究几何问题的方法.这种方法现在不仅广泛应用于数学的各个分支以及物理学、技术科学中的各个邻域,而且数形结合的思想对数学乃至整个科学的发展具有极其深远的影响.随着数学的发展,数和形两个方面都有了很大发展,数形结合的思想也就在更高的层次上开出绚丽的鲜花,结出丰硕的果实,代数拓扑就是其中之一.

拓扑学研究拓扑空间在同胚映射下不变的性质,即拓扑性质.所谓一个空间就是在一个集合中给定某种结构,拓扑空间是具有这样的结构的集合,使得能在其中定义映射的连续性,也就是要指明集合中什么样的子集是“开集”.把欧氏平面上开集所具有的性质抽象出来,据此规定一般集合上的开集,这个结构就是集合上的拓扑结构,也可以说拓扑结构就是开集结构——指明一个集合中的全部开集.具有拓扑结构的集合就是

① 原载自:王敬庚.直观拓扑[M].北京:北京师范大学出版社,2001:128-152.





一个拓扑空间.通常的欧氏空间就是一个拓扑空间.对于抽象的拓扑空间的讨论属于点集拓扑学(或一般拓扑学)的内容,点集拓扑中使用的是集合论的方法.我们只在欧氏空间中来研究图形的拓扑性质.圆盘、圆周都是欧氏平面上的图形,球和环面,圆柱面和麦比乌斯带都是三维欧氏空间中的图形,而射影平面、克莱因瓶都是四维欧氏空间中的图形.

代数拓扑是把拓扑空间中的图形与代数中的群和群的系列联系起来,也就是通过群和群的系列来描述它所联系的图形的几何结构,特别是图形中"洞"的排列情况.另外,若图形 X 联系群 $G(X)$,图形 Y 联系群 $G(Y)$,则从 X 到 Y 的连续映射 $f: X \to Y$ 也联系于群 $G(X)$ 到 $G(Y)$ 的同态 $f^*: G(X) \to G(Y)$,从而可以通过代数的研究解决拓扑的问题.

我们要求联系于一个空间的群一定是该空间的一个拓扑不变量,换句话说,同胚的空间有同构的群.这样这个群就给出比较空间不同胚的一个方法:若两个空间联系的群不同构,则它们一定不同胚.但是反过来不对,即若两个空间联系的群互相同构,一般地不能保证这两个空间互相同胚.(这一点与解析几何中曲线与方程的联系不同)

把群联系于空间的方法不止一种,我们这里介绍两种:同伦和同调,建立空间的基本群和同调群.无论哪种情况,目的都是相同的,即让代数的群结构反映空间的拓扑的和几何的结构.一旦这个群定义了并建立起它的性质,许多漂亮的几何定理便可用代数的论述加以证明.代数拓扑的威力就在于它是用代数方法解决拓扑的和几何的问题.

现在先简单地说一说同伦和同调的基本思想.

考虑曲线积分

$$\int_c p\mathrm{d}x + q\mathrm{d}y$$

此处 $p = p(x,y)$, $q = q(x,y)$ 是两个变量的连续函数,它们的偏导数连续且满足$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$.若分别沿着如图15.1所示的平环上的三条曲线 c_1, c_2, c_3 积分,则有

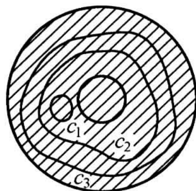


图 15.1

$$\int_{c_1} p\mathrm{d}x + q\mathrm{d}y = 0$$

及

$$\int_{c_2} p\mathrm{d}x + q\mathrm{d}y = \int_{c_3} p\mathrm{d}x + q\mathrm{d}y$$

这说明对于上述积分而言,c_1"等价"于常值道路,c_2 与 c_3"等价".如何描述两个闭道路是"等价"的?至少有下述两种方法.第一种方法:我们可以认为 c_2 和 c_3 是等价的,因为这两条道路中每一条可在平环内连续地变成另一条,这就是同伦理论的基本思想,并且说 c_2 与 c_3 同伦.c_1 同伦于常值道路(或平凡道路),

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

因为它可以在平环内连续地收缩到一点. c_2 与 c_1 不是同伦的道路,因为 c_2 不能越过它所包围的"洞"连续变成 c_1. 同理 c_3 与 c_1 也不同伦. 也可用另一种方法说 c_2 与 c_3 是等价的,因为 c_2 和 $-c_3$ 在平环中共同形成它们所包围的区域的边界,这是同调理论的基本思想,并且说 c_2 和 c_3 是同调的道路,说 c_1 同调于零,因为它是平环中它所围成的区域的完全边界. c_1 既不同调于 c_2,也不同调于 c_3.

同伦和同调的思想,是代数拓扑的创始人法国数学家庞加莱在1895年的一篇论文中引进的. 庞加莱和德国数学家希尔伯特齐名,同是他们那个时代的领先的数学家. 庞加莱的几何思想简直是惊人的. 他在微分方程(他的原来专长)、复变函数、代数、代数几何、天体力学、数学物理、天文学和拓扑学方面都作出了重要贡献. 他在数学方面写出三十本书和五百多篇论文,庞加莱的数学著作的数量仅次于欧拉. 此外,他在科学普及和数学哲学方面也是第一流的作家.

本章将尽可能直观地描述建立在同伦基础上的基本群以及建立在同调基础上的同调群.

2 道路的同伦类

本节介绍空间的道路、道路的同伦、道路的乘积、道路的同伦类等概念,为建立基本群做准备.

如图 15.2 所示,空间(这里指图形) X 中的一条以 x_0 为起点, x_1 为终点的道路 h,是指一个连续映射 $h:[0,1]\to X$,满足 $x_0=h(0)$, $x_1=h(1)$. 为了直观起见,在不引起混淆时,我们就把这个连续映射 h 的象集 $h([0,1])$—— 即单位线段 $[0,1]$ 在 h 下的连续象叫做道路 h(严格讲, h 和它的象集 $h[0,1]$ 是不等同的,因为存在不同的 h,但它的象集相同. 请读者自己举出这样的例子).

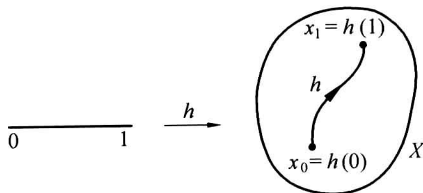


图 15.2

我们让一条道路在图 X 中作保持端点 x_0 和 x_1 不动的连续形变,在图 15.3(a) 上道路 h_1 形变为另一条道路 h_2 的过渡位置用细线描出. 并不是每一个图形上有相同端点 x_0 及 x_1 的任意两条道路,都可以由一条连续地变形为另一条. 例如在图 15.3(b) 表示的平环上,道路 h_1 和 h_2 之间能连续地互相形变,而无论 h_1 还是 h_2 都不能连续地形变为道路 h_3,尽管它们有相同的端点 x_0 及 x_1,

原因是因为 h_1 与 h_3 之间横着一个不可跨越的"洞".

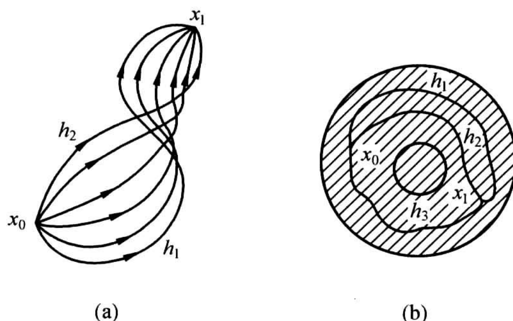


图 15.3

如果 X 中的两条道路 h_1 和 h_2,它们有共同的端点,并能经过连续形变由 h_1 变成 h_2,我们就称道路 h_1 同伦于道路 h_2,记为 $h_1 \sim h_2$.显然同伦是互相的,即若 $h_1 \sim h_2$,则 $h_2 \sim h_1$.

在圆盘内,任意两条有公共端点的道路都是互相同伦的.

当第二条道路 h_2 的起点恰好是第一条道路 h_1 的终点时,先沿第一条道路 h_1 再沿第二条道路 h_2,这样得到的道路称做道路 h_1 与 h_2 的乘积,记为 h_1h_2,它以第一条道路的起点为起点,第二条道路的终点为终点(图 15.4).

由道路乘积的这个定义看出,所谓两条道路相乘就是把两条道路接起来.因此并不是任意两条道路都能相乘,要能相乘必须第一条道路的终点恰为第二条道路的起点,因为只有这时,两条道路才能接起来.

若 $h_1 \sim h'_1, h_2 \sim h'_2$ 且 h_1 的终点恰为 h_2 的起点,则 $h_1h_2 \sim h'_1h'_2$(图 15.5).

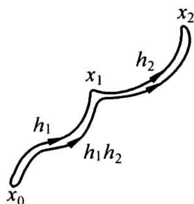


图 15.4

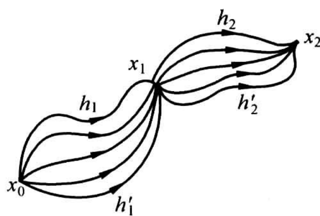


图 15.5

在有相同端点的所有道路的集合中,同伦关系是一个等价关系(即满足自反性、对称性和传递性).我们把所有互相同伦的道路归成一类,不同伦的道路归在不同的类,每一类叫做道路的一个同伦类或简称一个道路类,道路 h 所在的道路类记为[h].一个道路类中的每一条道路都称为这个道路类的一个

代表.

对于道路类也可以定义乘法.道路类$[h_1]$与$[h_2]$的乘积,记为$[h_1]\cdot[h_2]$,它是第一个类中的代表 h_1 与第二个类中的代表 h_2 的乘积道路 h_1h_2 所在同伦类,即

$$[h_1]\cdot[h_2]=[h_1h_2]$$

当然必须 h_1 的终点恰为 h_2 的起点才行.如果我们在这两类中分别另取代表 h'_1 和 h'_2,由前面的结果,所得乘积 $h'_1h'_2$ 是与 h_1h_2 同伦的,也就是它们属于同一个道路类,即$[h'_1h'_2]=[h_1h_2]$.所以两个类的乘积的定义为两个类所确定,与所选的代表无关.

现在来看道路的几种特殊情形.一条道路退化成一点 x_0,我们称之为常值道路(常用 c_{x_0} 表示).与道路 h 方向相反的道路,称为 h 的逆道路,记为 h^{-1}(图 15.6).起点和终点重合(都是 x_0)的道路,称为以 x_0 为基点的闭道路(也简称闭路或圈),如图 15.6 中的 α.闭路给我们提供了一个方便,即具有相同基点的任意两条闭路都可以相乘.

在空间 X 中,取定一点 x_0,在以 x_0 为基点的全体闭道路的集合中,按照同伦关系将它们分成互不相交的同伦类,把全体同伦类的集合记为 $\pi_1(X,x_0)$.

如果 X 是一圆盘,如图 15.7(a) 所示,任意取定一点 x_0,那么 X 中任意一条以 x_0 为基点的闭道路 α,都能连续形变到退化成一点 x_0 的道路 c_{x_0}.因此圆盘上每一条以 x_0 为基点的闭道路都与常值道路 c_{x_0} 属于同一个同伦类,即以 x_0 为基点的闭道路的同伦类只有一个$[c_{x_0}]$.这就是说当 X 是圆盘时,集合 $\pi_1(X,x_0)$ 只包含唯一的一个元素$[c_{x_0}]$.如界 X 是一平环,如图 15.7(b) 所示,任意取定一点 x_0,考察以 x_0 为基点的闭道路 α 和β.因为 α 不包围平环的洞,所以能收缩为一点 x_0,即 α 与常值道路 c_{x_0} 属于一个同伦类$[c_{x_0}]$.而 β 却因为它包围平环的洞,不能收缩成一点 x_0,因此 β 与 c_{x_0} 不属于同一类.这就是说当 X 是平环时,集合 $\pi_1(X,x_0)$ 中除了包含$[c_{x_0}]$ 以外还包含别的元素.

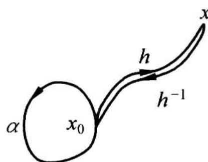


图 15.6

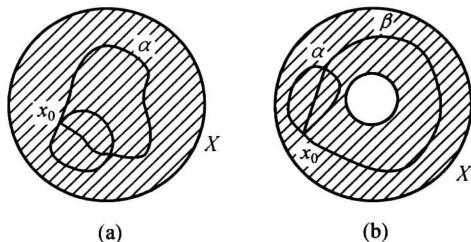
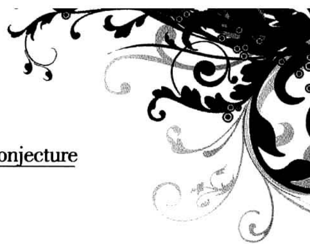


图 15.7



3 基本群

空间 X 中取定一点 x_0，以 x_0 为基点的闭道路的全体同伦类的集合记为 $\pi_1(X,x_0)$. 在上一节，我们已经对道路的同伦类定义了乘法，也就是在集合 $\pi_1(X,x_0)$ 中定义了一个乘法运算. 我们说对于这个乘法运算，集合 $\pi_1(X,x_0)$ 形成一个群. 按照集合形成群的要求，必须乘法满足结合律，有单位元，每一个元有逆元. 对此我们简单地加以说明.

首先考察以 x_0 为基点的常值闭道路 c_{x_0} 所在的同伦类 $[c_{x_0}]$. 对于 $\pi_1(X,x_0)$ 中的任一元素 $[\alpha]$，此处 α 是 X 中任意一条以 x_0 为基点的闭道路，由定义 $[c_{x_0}]\cdot[\alpha]=[c_{x_0}\alpha]$. $c_{x_0}\alpha$ 表示将闭路 α 的起点接在退化成一点 x_0 的常值道路 c_{x_0} 的终点 x_0 上(图 15.8)，可以认为它就是 α，即 $[c_{x_0}\alpha]=[\alpha]$. 于是 $[c_{x_0}]\cdot[\alpha]=[\alpha]$. 同理 $[\alpha]\cdot[c_{x_0}]=[\alpha]$. 所以 $[c_{x_0}]$ 对乘法 · 是单位元.

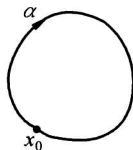


图 15.8

其次，对于 $\pi_1(X,x_0)$ 任一元素 $[\alpha]$，此处 α 是 X 中以 x_0 为基点的任意一条闭道路，考察 α 与 α 的逆道路 α^{-1}(即与 α 方向相反的闭道路) 的乘积 $\alpha\alpha^{-1}$. 我们注意到闭道路 $\alpha\alpha^{-1}$ 可以连续地收缩成一点 x_0(图 15.9)，于是 $\alpha\alpha^{-1}\sim c_{x_0}$，即 $[\alpha]\cdot[\alpha^{-1}]=[c_{x_0}]$. 同理 $[\alpha^{-1}]\cdot[\alpha]=[c_{x_0}]$. 我们把 $[\alpha^{-1}]$ 记为 $[\alpha]^{-1}$，它就是 $[\alpha]$ 的逆元. 这就是说 $\pi_1(X,x_0)$ 每一个元素 $[\alpha]$ 都有逆元 $[\alpha]^{-1}$.

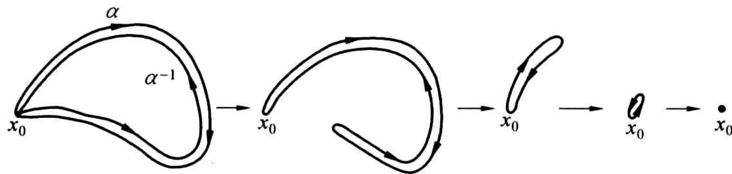


图 15.9

最后，对于 X 中以 x_0 为基点的任意三条闭道路，α,β,γ，乘积道路 $(\alpha\beta)\gamma$ 与 $\alpha(\beta\gamma)$ 可以认为是相同的闭道路，因此有 $([\alpha]\cdot[\beta])\cdot[\gamma]=[\alpha]\cdot([\beta]\cdot[\gamma])$，即乘法满足结合律.

所以，集合 $\pi_1(X,x_0)$ 对于道路类的乘法形成一个群. 这个群就称为空间 X 的以 x_0 为基点的基本群，仍用 $\pi_1(X,x_0)$ 表示. 基本群也称第一同伦群或庞加莱群.

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

在我们定义基本群时，是对于某个任意选取的基点 x_0 来进行的. 自然会问：选取不同的基点，对基本群有没有影响？可以证明若空间 X 是道路连通的（指空间中任何两点都有道路相连），则分别以任意两个不同的点 x_0 和 x_1 为基点的基本群是同构的，即 $\pi_1(X,x_0) \approx \pi_1(X,x_1)$. 因此对于道路连通的图形可以简单地说图形 X 的基本群，记为 $\pi_1(X)$，而不必特别指明基点. 如果图形分成互不连通的几块，则基点取在不同块上所作的基本群可以不同构，不过我们在这里讨论的图形都是道路连通的.

进一步我们指出：若空间 X 和 Y 同胚，则它们的基本群 $\pi_1(X)$ 与 $\pi_1(Y)$ 同构，即作为代数结构的基本群是空间 X 的拓扑不变量. 这样，如果两个空间的基本群不同构，则它们一定是不同胚的. 发现并研究基本群这个不变量的功绩归于庞加莱.

现在我们根据定义，通过直观的分析，找出一些常见图形的基本群.

圆盘的基本群. 由上一节讨论得到，圆盘 X 中以任一点 x_0 为基点的闭道路的同伦类只有一个，即$[c_{x_0}]$也就是 $\pi_1(X,x_0)$ 中只包含一个元素，即单位元，我们称这样的群为平凡群，于是圆盘的基本群为平凡群，记为 $\pi_1(X,x_0)=\{0\}$.

球面的基本群. 因为球面的表面上没有“洞”，因此球面上以任一点 x_0 为基点的闭道路都可以在球面上连续收缩到点 x_0(图 15.10). 因此任一条闭道路都同伦与常值闭道路 c_{x_0}，所以球面的基本群也是平凡群.

如果一个图形是道路连通的，且基本群为平凡群，则称这个图形是单连通的. 换句话说，如果一个连成一片的图形内任意一条闭道路都可以收缩成一点，就称这个图形是单连通的. 圆盘和球面是单连通的. 此外，直线、平面、线段、正方形片也是单连通的，(实心) 球体、(实心) 凸多面体以及凸多面体的表面也都是单连通的.

圆周 S^1 的基本群. 任取一点 x_0，以 x_0 为基点的任一条闭道路，若在某一段弧上先以一个方向通过一次，然后紧接着折回来以相反的方向再通过一次(图 15.11)，则可以通过连续形变将这个“皱折” 从道路中去掉(形变的过程可参看图 15.9)，不妨把这个小手术叫做“熨平”，经过熨平的闭道路和原来的闭道路是同伦的. 因此只需讨论熨平以后的闭道路有多少同伦类. 环绕圆周圈数不同的闭道路是不同伦的，因为中间隔着洞拉不过来；方向不同的闭道路也是不同伦的. 只有绕圈的方向相同且圈数也相同的两条闭道路才同伦. 我们指定依逆时针方向绕的圈数记为正数，依顺时针方向绕的圈数记为负数. 于是熨平以后的闭道路，按方向和圈数相同归在一类，共可分可数无限多类. 每一类对应一个整数：常值道路所在的类对应于0；依逆时针方向绕 n 圈(n 是正整数) 的闭道路所在的类对应于正整数 n；依顺时针方向绕 n 圈的闭道路所在的类对应于负整数 ——n，而且类的乘法对应于整数的加法. 因此圆周的基本群和整数加群同



构,记为 $\pi_1(S^1)\approx Z$.要知道,这里只是一种直观的描述,要给出 $\pi_1(S^1)\approx Z$ 的数学的严格证明并非轻而易举.

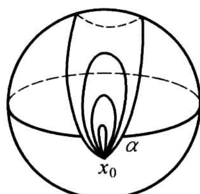


图 15.10

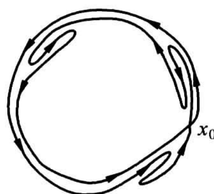


图 15.11

寻求平环及圆柱面的基本群,经过适当处理都可以化成求圆的基本群的问题,它们的基本群也都同构于整数加群 Z.

为了介绍基本群的一个应用,需要引进如下的概念.

我们对于每一个道路连通空间定义了基本群以后,进一步对于两个空间之间的连续映射,可以构造出它们的基本群之间的一个同态.(所谓两个群之间的同态,是指两个群的元素间的一个对应且保持运算)

设 $f:X\to Y$ 是空间 X 到 Y 的一个连续映射.取 x_0 为 X 中的基点,$f(x_0)=y_0$ 为 Y 中的基点.于是 X 中以 x_0 为基点的一条闭道路 α 在 f 之下的象就是 Y 中以 y_0 为基点的一条闭道路,记为 $f\alpha$,称为 α 的象道路.而且 X 中同伦的两条道路,在 Y 中的象道路也同伦.于是由 X 中的一个同伦类$[\alpha]$就得到 Y 中的一个同伦类$[f\alpha]$.这是从基本群 $\pi_1(X,x_0)$ 到 $\pi_1(Y,y_0)$ 的一个对应.把它记为 f_*.而且 f_* 把 $\pi_1(X,x_0)$ 中的$[\alpha]\cdot[\beta]$送到 $\pi_1(Y,y_0)$ 中的$[f\alpha]\cdot[f\beta]$,即保持运算.所以 f_* 是基本群之间的一个同态,称 f_* 为 f 的诱导同态.由此又可得到复合映射的诱导同态等于诱导同态的复合,即对于 $X\xrightarrow{f}Y\xrightarrow{g}Z$ 有 $(gf)_*=g_*f_*$.

4 同调群的直观描述

我们称某个曲面上的有向闭曲线或某些这种曲线的总和为这个曲面上的闭路,如图 15.12 所示环面上的定向经圆或定向纬圆或它们的总和都是闭路,曲面上的闭路可以分为两类,一类能界住某些区域(也就是作为某些区域的边界),另一类不能界住任何区域.这一概念,初看起来似乎是非常简单明显的(例如,平面上的圆周或其他任意一条自身不相交的闭曲线都界住它内部的区域),但实际上要求明确化时,就不是那么简单明显了.对于曲面上某一块区域 U 的边界曲线 z 按照如下方法选取方向:在曲面块 U 的法向正向一侧,面对这一

侧,沿着曲线 z 移动时,总使 U 在曲线的左边,就取这个移动的方向为 z 的方向.这时我们就把空间曲线 z 称为区域 U 的边缘(或者说 z 界住 U).如果把闭路 z 的方向反过来,则记为 $-z$.闭路界住某个区域并不像初看时那么简单,例如图 15.13 画着平面上的两个闭路 z_1, z_2,每个闭路界住它的内部区域(画有细线条),这两个闭路的和用 z_1+z_2 表示,它所界住的区域是由小块Ⅰ取两次,小块Ⅱ和Ⅲ各取一次组成,即闭路 z_1+z_2 界住的区域 $U=2\cdot \text{Ⅰ}+1\cdot \text{Ⅱ}+1\cdot \text{Ⅲ}$.这就是说闭路所界住的区域有时必须取曲面 X 的某些小曲面块带有确定的系数的和.图 15.12 所示环面上的闭路 z 界住一个区域,而环面上的经圆或纬圆或把两者合起来所成的闭路都不能界住任何区域.图 15.14 所示双环面上的闭路 z 界住画有细线条的区域.

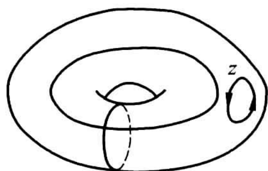


图 15.12

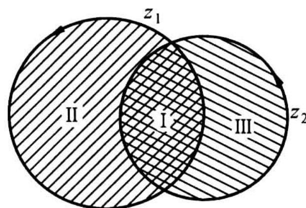


图 15.13

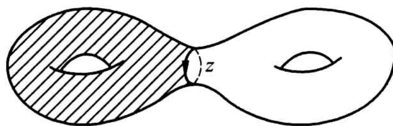
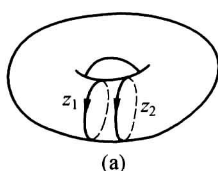


图 15.14

我们把界住某个区域的闭路称为“非本质的”,而把不界住任何区域的闭路称为“本质的”闭路.如果两个闭路 z_1 和闭路 z_2 一起共同形成一个区域的边缘,我们就称闭路 z_1 与 z_2 是“非本质地不同”的,也称闭路 z_1 与闭路 z_2 彼此同调.例如图 15.15(a)中环面上的两个具有所示方向的经圆 z_1 与 z_2 就是同调的,因为图 15.15(b)中 z_1 与 $-z_2$ 的和界住画有细线的区域.

为了认清曲面包含“洞”的情况,我们感兴趣的是曲面上有多少彼此本质地不同的本质的闭路,也就是有多少彼此不同调的非边缘的闭路?显然环面上的上述这种闭路多于球面,而双环面上的这种闭路又多于环面.图 15.16 所示双环面上的四条闭路.两条定向经圆 m_1, m_2 和两条定向纬圆 p_1, p_2 都是本质地不同的即互不同调,又都不是某个区域的边缘.



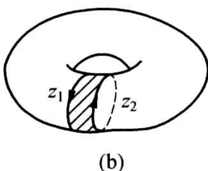


图 15.15

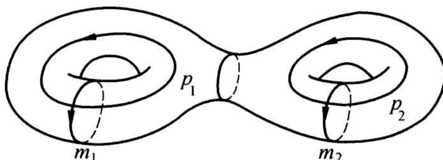


图 15.16

闭路的同调关系是一个等价关系(即是自反的、对称的和传递的).将给定曲面的所有闭路按同调关系进行分类,把互相同调的闭路归入一类,互相不同调的归入不同的类,每一类称为一个同调类.例如球面上所有闭路都是某个区域的边缘.如果一个闭路恰是某个区域的边缘,我们也称该闭路同调于零.因此球面上的所有闭路都是互相同调的,它们都同调于零.在环面上存在无限多个同调类,但它们都可以用两个基本同调类,即经圆类和纬圆类表示.例如,按纬的方向环绕环面三次,按经的方向环绕环面两次的所有闭路(图 15.17)形成的同调类可以表示为 $2m+3p$,这里 m 表示经圆类,p 表示纬圆类.一般说,环面上的任一同调类具有形式 $\alpha m+\beta p$,这里 α,β 是整数.而且同调类之间可以进行加减(就像二元一次齐次多项式那样进行加减),所以环面上全体同调类的集合形成一个群.这个群是具有两个生成元的自由交换群,同构于两个整数加群的直和,记为 $Z\oplus Z$.在双环面上有四个生成元 m_1,m_2,p_1,p_2(图 15.16),每一同调类具有形式 $\alpha_1m_1+\alpha_2m_2+\beta_1p_1+\beta_2p_2$,全体同调类的集合组成一个四个生成元的自由交换群.我们把一个曲面上的全体同调类的集合形成的群叫做这个曲面的同调群.球面的(一维)同调群只包含一个元素是平凡群.环面和双环面的(一维)同调群是分别有两个生成元和四个生成元的自由交换群.

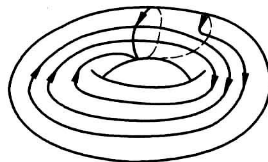


图 15.17

要在一个曲面上构造同调群,也就是要找出该曲面上闭路的所有同调类.闭路 z_1 和 z_2 同调当且仅当 z_1 和 $-z_2$ 一起共同组成一个区域的边缘,即 $z_1+(-z_2)=z_1-z_2$ 是一个边缘.特别地当一个闭路恰是某区域的边缘时,说它同调于零.现在我们用代数中群的有关术语来描述同调群.

已知一个交换群 G(运算用加法表示)和它的一个子群 H,我们可以利用 H 来将 G 中的元素分类.规定 G 中元素的一个等价关系:$a\sim b$ 当且仅当 $a-b\in H$(容易验证这恰是一个等价关系),按照这个等价关系可以将 G 中的元素分类.a 所在的等价类用$[a]$表示,定义一个加法$[a]+[b]=[a+b]$,则 G 中所有等价类的集合对于这个加法也构成一个群,称为群 G 模子群 H 的商群,记为

G/H.形象地说 G 中的一个元素“模掉”(减去)属于 H 的部分剩下的就是商群 G/H 的一个元素;G 中两个元素各模掉属于 H 的部分以后相等,那么它们就属于商群中的同一个元素.这样 H 中的所有元素都属于商群中的同一元素(单位元 0).例如,整数模 2 剩余类群是一个整数加群模偶数群 $\{2n \mid n \in \mathbf{Z}\}$ 所得商群,因为任一整数模去偶数以后或者是 0 或者是 1,所以,该商群只有两个元素[0]及[1].

回到我们的问题上来.首先注意到曲面上由所有闭路组成的集合对于加法组成一个交换群,暂且叫它闭路群.其中所有界住某个区域的闭路即边缘,也构成一个群(因为两个边缘的和或差仍然是一个边缘),暂且叫它边缘群,它是闭路群的一个子群.我们取闭路群模边缘群的商群,按商群的定义,两个闭路 z_1 和 z_2 属于商群中的同一个元素时,当且仅当它们的差 $z_1 - z_2$ 属于边缘群,即 $z_1 - z_2$ 恰好界住某个区域,也就是它们彼此同调.所以商群的元素就是同调类.因此曲面的同调群就是曲面上的闭路群模边缘群的商群,通俗地说就是把所有的边缘忽略不计.

如上考察的只是二维图形(曲面)上的一维的闭路(闭曲线),在拓扑学中要考察不同维数的图形和不同维数的闭路(零维、一维、二维、三维等等).如上讨论的只是一维同调群,记为 $H_1(X)$.同样的讨论可以得到其他各维的同调群 $H_0(X)$, $H_2(X)$, $H_3(X)$, $\cdots$, $H_n(X)$.

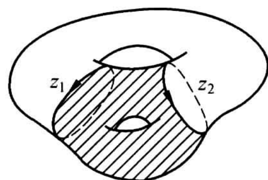


图 15.18

关于两个闭路同调的概念和同伦的概念之间是什么关系呢? 首先,若两个闭路 z_1 和 z_2 是同伦的,即由闭路 z_1 经过连续形变可得到 z_2,于是 z_1 在这个变形的过程中“扫过”的“轨迹”就是连接 z_1 和 z_2 的区域,z_1 和 z_2 就共同组成这个区域的边缘(图 15.15),因此得到 z_1 和 z_2 是同调的.但反过来,不一定对,也就是说两个同调的闭路不一定同伦,如图 15.18 所示曲面的闭路 z_1 和 z_2 是同调,因为 $z_1 - z_2$ 恰是用细线画出的区域的边缘,但 z_1 却不能在曲面上连续地形变到 z_2,因为它受到横在这两个闭路之间的“洞”的阻碍,即 z_1 和 z_2 不是同伦的.因此我们得到,两个闭路同伦是这两个闭路同调的充分条件但非必要条件.

最后我们指出,在任一空间中,各维闭曲线按同调分类的问题是极其困难的.为了避开这个困难,在单纯同调理论中,我们只研究这样一类特殊的空间.在这类空间中我们能够找到一个很自然的,更重要的是很容易处理的闭曲线类,用它可以去检测空间中的洞.这一类空间同胚于欧氏空间中的多面体,叫做可单纯剖分的空间.

5 闭链、边缘链和同调群

为了便于计算同调群,我们只研究一类特殊的空间,它可以看成是由一小块一小块我们所熟悉的空间很好地拼凑起来的,就是所谓可单纯剖分的空间.因为曲面是二维的,因此可以用三角形来拼成它们的模型,对于高维空间,就需要相当于二维三角形的"高维砖块"来砌成我们的建筑物了.

我们在欧氏空间中来讨论,我们再次申明讨论中对于一些概念我们不给出严格的数学定义,只是尽可能通过具体例子来说明.我们把由一点组成的图形叫 0 维单纯形,简称 0 维单形.闭线段叫做 1 维单形,三角形叫做 2 维单形,四面体叫做 3 维单形(图 15.19).每一单形中包含的比它自己的维数低的单形称为该单形的(真)面单形(简称面).我们把一个单形的顶点排列起来表示这个单形,例如图 15.19 中的单形分别记为 $\sigma^0 = < v_0 >$, $\sigma^1 = < v_0, v_1 >$, $\sigma^2 = < v_0, v_1, v_2 >$, $\sigma^3 = < v_0, v_1, v_2, v_3 >$.单形 $\sigma^3 = < v_0, v_1, v_2, v_3 >$ 的 0 维面(即顶点)有 $< v_0 >$, $< v_1 >$, $< v_2 >$, $< v_3 >$;1 维面(即棱)有 $< v_0, v_1 >$, $< v_0, v_2 >$, $< v_0, v_3 >$, $< v_1, v_2 >$, $< v_1, v_3 >$, $< v_2, v_3 >$;2 维面有 $< v_0, v_1, v_2 >$, $< v_0, v_1, v_3 >$, $< v_0, v_2, v_3 >$, $< v_1, v_2, v_3 >$.我们所说的两个单形"很好地"连接在一起(或说它们规则相处)是指它们要么不相交,如果相交,那么公共部分是它们的公共面单形(图 15.20).

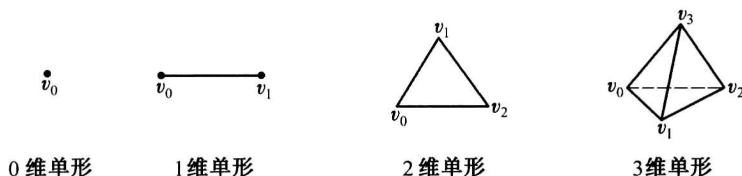


图 15.19

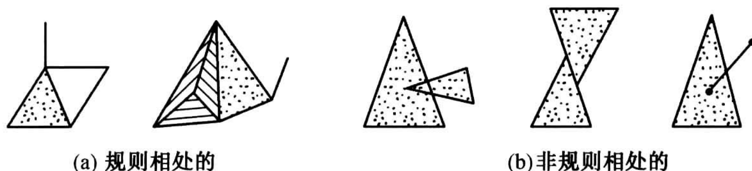


图 15.20

欧氏空间中由有限多个单形组成的如下集合称为一个单纯复形(简称复形),其中所有的单形都是规则相处的,并且只要一个单形属于这个集合,那么它的所有面也属于这个集合.我们用字母 K 表示复形,它是欧氏空间中的一组

单形.如果把每个单形看成欧氏空间中的一个点集,那么 K 中单形的并集就是欧氏空间的一个子集,可以成为一个拓扑空间(欧氏空间的子空间).我们把这个拓扑空间叫做联系于复形 K 的一个多面体,记为 $|K|$.一个空间 X 叫做可单纯剖分的,如果存在一个复形 K,使得 X 同胚于 K 所联系的多面体 $|K|$,复形 K 就称为空间 X 的一个单纯剖分,例如四面体的表面上的那些单形组成的复形是球面的一个单纯剖分.

要求空间可以单纯剖分是一个很强的要求,然而很多重要的空间都是可剖分的,例如所有的闭曲面都是可单纯剖分的,单纯剖分也叫三角剖分,单纯剖分不是唯一的.

现在我们回到同调群的计算上来.

以环面为例,我们将限于考虑环面的一个选定的单纯剖分 K.考虑这个剖分之下的定向多边形曲线,方向用箭头标在多边形的棱上.若一条棱的顶点为 v,w,则用记号 $<v,w>$ 表示具有从 v 到 w 的方向的这条棱.类似地,若 u,v,w 是 K 的一个三角形的顶点,则 $<u,v,w>$ 表示按顶点次序 u,v,w 定向的这个二维单形.因此有

$$<u,v,w> = <v,w,u> = <w,u,v>$$

即只要保持三角形顶点的环绕方向不变,从哪个顶点开始写是没有关系的.序向的改变用一个负号来表示,即

$$<w,v> = -<v,w>,\ <v,u,w> = -<u,v,w>$$

定向棱 $<v,w>$ 的边缘定义为

$$\partial<v,w> = <w> - <v>$$

定向三角形的边缘定义为

$$\partial<u,v,w> = <v,w> + <w,u> + <u,v>$$

也就是定向单形 $<u,v,w>$ 的边缘是它的棱的和,每条棱的序向是由三角形的序向诱导来的(图 15.21).

0 维单形边缘定义为零.

现在我们来研究如何判别曲面上的一条定向曲线是否为闭的,进一步是否为曲面上一小块曲面的边界？图 15.22 中的定向曲线 A 看做是各定向棱之和

$$A = <u,v> + <v,w> + <w,x> + <x,y> + <y,u>$$

我们定义 A 的边缘为它的各定向棱的边缘之和

$$\partial A = \partial<u,v> + \partial<v,w> + \partial<w,x> + \partial<x,y> + \partial<y,u>$$

分别计算各定向棱的边缘再加起来得

$$\partial A = <v> - <u> + <w> - <v> + <x> - <w> + <y> - <x> + <u> - <y> = 0$$

于是得知曲线 A 没有边缘,即 A 是闭的.这就是判别一条曲线是闭的方法.用



这个方法可知图 15.22 中的定向曲线 B

$$B=<a,b>+<b,c>+<c,d>+<d,e>+<e,a>$$

也是闭的. 而且从图我们还看到闭曲线 B 恰好包住了三个三角形. 如果这三个三角形的定向如图 15.22 所示,把由这三个定向三角形拼在一起所得的一小块曲面 P 用下式表示

$$P=<e,a,b>+<e,b,c>+<e,c,d>$$

计算这一小块曲面 P 的边缘

$$\begin{aligned}\partial P=&\partial<e,a,b>+\partial<e,b,c>+\partial<e,c,d>=\\&<a,b>+<b,e>+<e,a>+<b,c>+<c,e>+\\&<e,b>+<c,d>+<d,e>+<e,c>=\\&<a,b>+<b,e>+<e,a>+<b,c>+<c,e>-\\&<b,e>+<c,d>+<d,e>-<c,e>=\\&<a,b>+<b,c>+<c,d>+<d,e>+<e,a>=B\end{aligned}$$

这就是判别曲线 B 是曲面上某一小块曲面的边缘的方法.

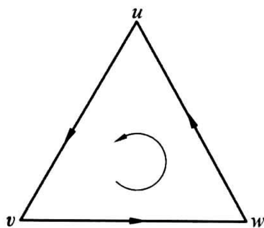


图 15.21

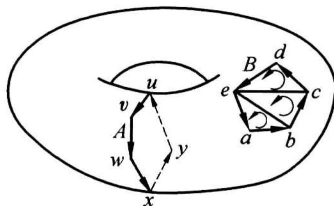


图 15.22

现在考虑一般情形,K 的定向棱的以整数为系数的一个线性组合

$$z=\lambda_1<u_1,v_1>+\cdots+\lambda_k<u_k,v_k>$$

如果它的边缘是零,即

$$\partial\lambda_1<u_1,v_1>+\cdots+\partial\lambda_k<u_k,v_k>=0$$

我们就叫 z 为 K 的一个 1 维闭链. 对于 K 的任意两个 1 维闭链 $\sum\lambda_i<u_i,v_i>$ 及 $\sum\mu_i<u_i,v_i>$,定义加法

$$\sum\lambda_i<u_i,v_i>+\sum\mu_i<u_i,v_i>=\sum(\lambda_i+\mu_i)<u_i,v_i>$$

则 K 的全体 1 维闭链在这个加法下构成一个交换群,叫做 1 维闭链群,记为 $Z_1(K)$.

K 的一个定向的简单闭多边形曲线看做是它的定向棱之和时,是特别简单的 1 维闭链,我们把它称之为基本 1 维闭链,可以证明 1 维闭链群 $Z_1(K)$ 可以由这些基本 1 维闭链生成.

根据前面对闭曲线 B 的考察,我们得到:如果对于一个 1 维闭链能找到定向三角形的一个整系数的线性组合,使得它的边缘恰是已知的这个闭链,我们就说这个 1 维闭链是一个 1 维边缘闭链,或简称 1 维边缘链.全体 1 维边缘链组成 1 维闭链群的一个子群,称为 1 维边缘链群,记为 $B_1(K)$.我们感兴趣的是包围一个“洞”的闭链,这种闭链不是边缘链.因此我们要的是非边缘的闭链,而把边缘链忽略掉.为此我们作 1 维闭链群模 1 维边缘链群的商群,记为叫做 K 的 1 维同调群

$$H_1(K) = Z_1(K)/B_1(K)$$

根据同调群的这个定义,若两个闭链的差是一个边缘链,则它们代表 $H_1(K)$的同一个元素,这时这两个闭链叫做同调的.例如图 15.23 的闭链 z_1 和 z_2 是同调的,因为 $z_1 - z_2$ 是它们之间的管形区域(它的三角形定向如图)的边缘.

计算可得环面的 1 维同调群 $H_1(K) \approx Z \oplus Z$(表示同构于两个整数加群的直和).它的两个生成元可以用基本闭链 z_1 和 z_2 来代表.这里 z_1 是一个经圆,z_2 是一个纬圆.任何其他的 1 维闭链必定同调于这两个闭链的线性组合.例如图 15.24 中对角线表示的闭链 z 同调于 $z_1 + z_2$,这是因为从图上可以看出 $z_1 + z_2 - z$ 是半个环面的边缘.

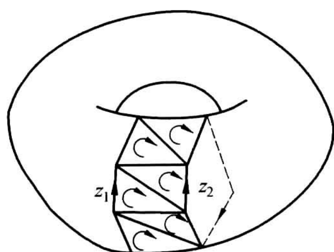


图 15.23

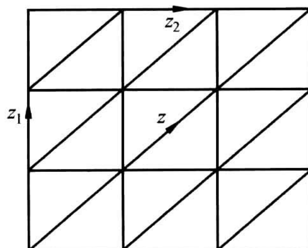


图 15.24

环面的其他各维同调群是 $H_0(K) \approx Z$, $H_2(K) \approx Z$, $n \geqslant 3$ 时, $H_n(K) \approx \{0\}$(因为 K 中没有 2 维以上的单形).

现在来计算圆盘的同调群.首先给出圆盘的一个单纯剖分,它是 2 维单形 $< v_0, v_1, v_2 >$ 及它的所有面单形组成的复形 K,即 K 中包含 0 维单形 $< v_0 >$, $< v_1 >$, $< v_2 >$,定向 1 维单形 $< v_0, v_1 >$, $< v_1, v_2 >$, $< v_2, v_0 >$ 及定向 2 维单形 $< v_0, v_1, v_2 >$(图 15.25).

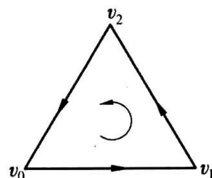


图 15.25

K 的 0 维链是 K 的 0 锥单形的下列形式和

$$c_0 = g_0 < v_0 > + g_1 < v_1 > + g_2 < v_2 >$$



这里 g_0, g_1, g_2 是整数. 因为每一个 0 维单形的边缘定义为零,所以任一个 0 维链的边缘是零,即每一个 0 维链皆为 0 维闭链. 求 K 的 0 维同调群,也就是求 0 维闭链的同调类的集合. 注意到

$$g_0 < v_0 > + g_1 < v_1 > + g_2 < v_2 > =$$
$$\partial(g_1 < v_0, v_1 > + g_2 < v_0, v_2 >) + (g_0 + g_1 + g_2) < v_0 >$$

即 $c_0 - t < v_0 >$ 恰是一个 1 维链的边缘,此处整数 t 是 c_0 中的 0 维单形的系数和,于是 0 维链 c_0 同调于 $t < v_0 >$. 因此每一个 0 维同调类有一个代表 $t < v_0 >$,并且当 $t_1 \neq t_2$ 时, $t_1 < v_0 >$ 与 $t_2 < v_0 >$ 不同调,这是因为

$$t_1 < v_0 > - t_2 < v_0 > = (t_1 - t_2) < v_0 >$$

不是任何 1 维链的边缘. 又由于 t 可以取遍所有整数值,所以 0 维同调群同构于整数加群,即 $H_0(K) \approx Z$.

再来计算 K 的 1 维和 2 维同调群. K 的 1 维链形如

$$c_1 = h_0 < v_0, v_1 > + h_1 < v_1, v_2 > + h_2 < v_2, v_0 >$$

这里 h_0, h_1, h_2 是整数. 它的边缘

$$\partial c_1 = (h_2 - h_0) < v_0 > + (h_0 - h_1) < v_1 > + (h_1 - h_2) < v_2 > \quad ①$$

于是 c_1 是 1 维闭链,即$\partial c_1 = 0$ 当且仅当 $h_0 = h_1 = h_2$,因此 1 维闭链形如

$$z_1 = h < v_0, v_1 > + h < v_1, v_2 > + h < v_2, v_0 > \quad ②$$

此处 h 是任意整数. 于是得到 1 维闭链群 $Z_1(K) \approx Z$.

K 只有一个 2 维单形 $< v_0, v_1, v_2 >$,因此 2 维链形如 $h < v_0, v_1, v_2 >$,此处 h 是整数. 由于

$$\partial(h < v_0, v_1, v_2 >) = h(\partial < v_0, v_1, v_2 >) =$$
$$h < v_0, v_1 > + h < v_1, v_2 > + h < v_2, v_0 > \quad ③$$

因此 2 维链是闭链当且仅当 $h = 0$,即所有的 2 维闭链皆为 0,即 $Z_2(K) = \{0\}$,于是 $H_2(K) = \{0\}$.

由等式②及③得到,每一个 1 维闭链皆为 1 维边缘链,因此 $Z_1(K) = B_1(K)$,于是 $H_1(K) = \{0\}$.

因为 K 中没有 2 维以上的单形,所以 $n > 2$ 时, $H_n(K) = \{0\}$.

这样得到圆盘的各维同调群 $H_0(K) \approx Z, H_1(K) = \{0\}, H_2(K) = \{0\}, n > 2$ 时, $H_n(K) = \{0\}$.

上述证明 $H_n(K) \approx Z$ 的方法具有一般性,对于行意连通复形都可用. 所谓一个复形 K 是连通的,是指 K 中的任意两个 0 维单形都可以用一连串首尾相接的一维单形把它们"连接"起来. 对于任意连通复形 K 有 $H_0(K) \approx Z$,证明留作练习.

现将一些常见图形的各维同调群列表如下:

图形名称	$H_0(K)$	$H_1(K)$	$H_2(K)$	备　注
圆　盘 （即2维单形）	Z	$\{0\}$	$\{0\}$	n维单形，$n>2$ （$n=3$时是实心球体） $H_0(K)\approx Z$ $H_i(K)=\{0\},0<i\leqslant n$
平　　环	Z	Z	$\{0\}$	
麦比乌斯带	Z	Z	$\{0\}$	
球　面 （2维球面）	Z	$\{0\}$	Z	n维球面，$n>2$ $H_0(K)\approx Z$ $H_n(K)=Z$ $H_i(K)=\{0\},0<i<n$
环　　面	Z	$Z\oplus Z$	Z	
射影平面	Z	Z_2	$\{0\}$	Z_2是整数模2 同余类群
克莱因瓶	Z	$Z\oplus Z_2$	$\{0\}$	

虽然我们是对多面体的某一个选定的单纯剖分（即复形）来定义和计算其同调群的，但是经过复杂的步骤，我们能够证明同调群是拓扑不变量，与单纯剖分的选择无关.

根据定义，各维单纯同调群 $H_p(K)=Z_p(K)/B_p(K)$都是有限生成的交换群.因此都可以写成 $F\oplus T$ 的形状，其中 F 是有限生成的自由交换群，即有限多个整数加群 Z 的直和，T 是一个有限交换群. T 的元素是同调群中的有限阶元素，称为挠元素，反映了曲面的“扭曲”. F 的秩，也就是 Z 的直和因子的个数，叫 K 的 p 维贝蒂数，它表示多面体 $|K|$ 上“p 维洞”的个数.

切赫（Čech，Eduard，1893—1960），捷克数学家.

与基本群及高维同伦群相比，单纯同调群的计算方便些，但为此付出的代价是对研究的空间加以相当严格的限制，即只研究多面体这类空间.除了单纯同调理论以外也还有更一般的同调理论，它们研究更广泛的空间类，例如奇异同调论及切赫同调论，它们比单纯同调论更加抽象.





佩雷尔曼和俄罗斯拓扑学传统[1]

第十六章

苏维埃时代的俄罗斯数学发展里,一个最重要的现象是,俄罗斯拓扑学派的形成.这个方兴未艾的数学部门——拓扑学——的科学研究,开始于19世纪末年庞加莱及嘉当的著述.拓扑学的发生发展,以及其他各种新的数学分科,像一般集论、抽象代数、数理逻辑的发达,正是现代数学的特征.现在这个时代,开启了新的宝库,并重新估价了旧的数学.在苏联,这数学的新方向,得到了光辉的发展.俄罗斯拓扑学派,开始创立于20世纪20年代,但在目前已雄踞了世界的首位.

俄罗斯拓扑学派的重要创始人就是П.С.Александров.在苏联,他与夭逝的不世出的天才П.С.Урысон首先共同在一般拓扑空间领域里,开始工作.不久他的同门弟子А.Н.Тихонов,Н.Ь.Веденисов,В.В.Немыцкй,Л.А.Тумаркин等,加入了他们的研究.1926年奠定了莫斯科拓扑学集团.就这样产生了俄罗斯拓扑学派.那时只能叫做"莫斯科",因为我们现在的数学研究中心彼得格勒等处,还是"古典"方向占优势,在那个时期,对于这个研究领域,并没有发生了什么重要的指南作用.

俄罗斯拓扑学派开拓了许多的重要领域.这些里包含了"组合"拓扑学,并奠定了组合拓扑学与一般拓扑学空间论的关系.在20世纪20年代末,П.С.Александров的高足П.С.Понтрягин对拓扑学的各方面,作了重要的贡献.在谈俄罗斯拓扑学派的时候,是有彼得格勒在内的.

① 原载自:А.А.Марков等.拓扑学与描述集合论(三十年来的苏联数学)[M].杨宗磐,孙以丰,关肇直,译.北京:科学出版社,1955.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

20世纪40～50年代的俄罗斯拓扑学派是由П.С.Александров领导，并且有很多的学者，有成就地研究着各种问题.

在这篇文章，我想将这个学派的重要成就，试作一个简明的叙述.

1.拓扑学可以定义为所谓“拓扑空间”的理论.目前在数学里，拓扑空间的定义是

在一个任意的元(叫元作点)所构成的集里，同时考虑某些下集(称之为开集)，满足：

E_1 点全体所成集是开集.

E_2 任意开集的集的和，仍是开集.

E_3 任两个开集的交，仍是开集.

换句话说，要想从集 X 得到拓扑空间，必须在 X 的下集里，选出某些满足 E_1，E_2，E_3 的集才可以.X 的元叫做点，选出来的集叫做开集，选出来的集全体将 X 定义为空间.

按这个定义，空间 X，Y 一致的充要条件是 X，Y 有相同的点，还有相同的开集.上述空间的概念与1922年波兰学者Kuratowski所下的定义同值.这个概念还可以有各种不同的推广.А.Д.Александров[1]曾经作了下述有意义的推广：就是说不用 E_2，而代替之以更弱的条件

$E_2{}^*$ 不比可数多的开集的集的和，仍是开集.

其他公理仍旧.

А.Д.Александров的空间与测度的拓扑理论有密切关系，又与下边要提到的很多拓扑学问题有关.

我们要注意，由 $E_2{}^*$ 或由 E_2 可以得出，空集是开.理由是空集与开集的空集的和是一样.从 E_3 可以得出，任意的开集的不空有穷集的交是开集.

2.拓扑空间的定义有了之后，我们就可以作下列诸定义：“闭集”、“闭包”、“点的近傍”、“集的近傍”、“连续变换”、“拓扑变换”、“拓扑同值”、“拓扑性质”、“拓扑不变性”、“通空间”、“通集”、“空间成分”、“孤点”，以及关于这些概念的简单定理.我不预备在这里详述这些定义及定理，读者可以参阅豪斯道夫的*Mengenlehe*①或П.С.Александров的论拓扑学中的空间观念.

3.在拓扑学里，有几种运算具有重要的作用.它能从一个空间做出新的空间来，其中一个构成性的运算是作“下空间”，做法是这样的.

设 Y 是空间 X 的下集.定义空间 X 的开集与 Y 的交为 Y 的开集.很容易看出来，这样就可以把 Y 变成拓扑空间.这个空间，我们叫做用 Y 定义的 X 的下空间.若 L 是一个空间类，属于 L 的空间的任一下空间若仍属于 L 的时候，我

① 有俄译本：Ф.Хаусдорв，Теория Множесть.М.－Л.，ОНТИ(1937)，1-302.





们说 L 具遗传性.

4.另一个构成性的重要运算,是由空间系作“积”.这个运算是 A.H. Тихонов[2]所引进的.其定义如下:

不空集 Ξ 的每个元 ξ 对应一空间 X_ξ.考察系 $\langle X_\xi\rangle \xi\in\Xi$ 全体所作的 X,这里 X_ξ 是 X_ξ 的点,ξ 属于 Ξ.X 的开集的定义是:任选一个 Ξ 的元 α,固定之,又任取 X_α 的开集 A.用符号 $Z(\alpha,A)$表示 X 的 $\langle x_\xi\rangle \xi\in\Xi$ 而 $x_\alpha\in A$ 的全体.所有 Ξ 的 α,及所有的 X_α 的开集 A 所对应的 $Z(\alpha,A)$都含于 X.有穷个 $Z(\alpha,A)$的交叫做 X 的基.X 的开集是任意基集的和.很容易知道,这样定义的 X 的下集满足 E_1,E_2 及 E_3.因之 X 成了拓扑空间.这样做成的拓扑空间 X 叫做空间系 $\langle X_\xi\rangle \xi\in\Xi$ 的积.

若 L 是一个空间类,由属于 L 的空间所成的系做出的积,若仍属于 L 的时候,我们说 L 具乘性.

特别每个空间因子都等于一个空间 Y 的时候,我们叫做 Y 的 τ 幂,这里 τ 是这唯一因子所成系的势.这个名称是对的,因为除拓扑变换不计外,Y 的 τ 幂是用 Y 及基数 τ 定义的,用符号 Y^τ 表示它.

特别令 $Y=D$,D 是两孤点所成空间,于是对于任一基数 τ 得到空间 D^τ.空间 $D^{\aleph_0}$拓扑同值于嘉当的二进不连续集.D^τ 的成分全都是单独的一点.同时,对于无穷的 τ,D^τ 自稠,换句话说没有孤点.无穷 τ 所对应的 D^τ,叫做一般嘉当不连续集.

若 $Y=I$,I 是用通常开集所规定的直线上的单位线段,于是,得到 A.H. Тихонов 首先考虑的 I^τ.我们把它叫做 Тихонов 方体,因为 I^3 就是普通的立方体.Тихонов 方体是通集,这由 I 的通性及通空间的乘性很容易可以知道.

5.空间 X 的开集系 β 若能满足:所有 X 的开集都能由 B 的某些(任意个)集所成和得出的时候,我们叫 β 作 X 的基.任何空间都有基.例如取空间里所有的开集所成集系就是基.空间 X 的不同的基的势里有一个最小的.我们叫这最小数做空间 X 的重.

很容易看出:空间 X 的任一下空间的重不会大过 X 的重.因此得出:任予基数 τ,重$\leqslant\tau$ 的空间,产生具遗传性的类.

关于积空间的重有下面一个很重要的定理,这定理为彼得格勒数学家 М.Я.Перелвман①在 1942 年得到的.

积空间的重等于因子空间的重的和,但条件是要和为无穷,并且每个被加重大于 1.

由此特别可以得出 Тихонов 方体 I^τ(τ 无穷)的重是 τ(Тихонов[2]).n 维数

---

① Мощности в топологии.Л.,Рукописв(拓扑学里的势,稿本),1941:1-361.

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

空间是数直线的 n 幂，所以重是 $\aleph_0$.

6.空间 X 的下集 A 的闭包等于 X 的时候，我们说 A 稠于 X. 稠于空间 X 的集的势中的最小的叫做空间 X 的拟重，空间的拟重不会超过空间的重，或许较重会小.

7.要想做出方便有用的拓扑结构，必须引进某些用附加公理定义的空间. 选择这些附加公理，基本上决定于拓扑学及其他数学分支之间的，主要是与分析及几何之间的互相影响. 研究者设法沟通拓扑学及这些数学部门，于是建立起由这些数学分支看，非常有兴趣的空间的公理特征. 内在的纯逻辑动机也未尝没有. 就这样达到了现代的精炼的拓扑公理论，对于这个工作，苏联拓扑学者贡献很大. 我们现在就来考察这公理论及其有关的一般问题.

8.在拓扑公理论里，“分离公理”占很重要的位置. 分离公理同下列的诸概念有密切的关系.

设 A，B 是空间 X 的下集. 若 A 在 X 有近傍与 B 不相交，我们说 A 在 X 与 B 近傍分离，若 A，B 在 X 分别有近傍 U，V 互不相交，我们说 A，B 在 X 互相近傍分离. 若有实值连续函 f 定义于 X，在集 A 的所有点只取值 0，在集 B 的所有点只取 1，我们说 A，B 在 X 泛函分离. 这里以及以后说到定义在 X 的连续函都指的是，从空间 X 到用普通方法拓扑化的数直线内的连续变换.

9.我们所考虑的分离公理里最弱的那个公理是 А. Н. Колмогоров 所提出的，这公理是：

空间的任两一点集(由一个点所构成的集)中至少有一个在空间与另一个近傍分离.

换句话说，空间的任两不同点 x，y，或 x 有近傍不含 y 或 y 有近傍不含 x. 满足这个公理的空间叫做 Колмогоров 空间，这个空间类用符号 T_0 表示.

从一般的空间转换成 Колмогоров 空间，本质上并没有缩小了研讨的范围. 因为考察任意的空间时，经过适当的点的“斩齐”手续，可以归于考察某些 Колмогоров 空间的.

很容易看得出，T_0 是具遗传性及乘性的.

下面所举 Колмогоров 空间的例最简单，而同时有说一说的价值. 设空间 F 的点是数 0 及 1，开集是：Λ，$\{0\}$，$\{0,1\}$①. 于是集 $\{0\}$ 近傍分离于集 $\{1\}$，而集 $\{1\}$ 并不近傍分离于 $\{0\}$.

因为 T_0 具乘性. 所有的 F^{τ}，τ 是任意基数，属于 T_0. 又因为 T_0 具有遗传性，所以任一空间 F^{τ} 的所有下空间都属于 T_0. 我们立刻可以知道，T_0 类实质上尽于 F^{τ}.

---

① 此处及以后“Λ”表示空间，“$\{x\}$”表示由一个元 x 所成的集，“$\{x, y\}$”表示由元 x，y(不一定不相同)所构成的集，余类推.





10. 为了便于精确、简明地叙述 П.С. Александров 的重要定理,以及以后谈到的结果起见,我们引进下列术语:

若空间 X 拓扑同值于空间 Y 的某一下空间时,我们说 X 寝于 Y.

设 R 是具遗传性的空间类.若 U 属于 R,并且所有属于 R 的空间都属于 U,我们说 U 是 R 类的万有空间.

很显然地,研讨空间类时,若能建立某些空间的万有性,就是在这空间类的理论上前进了重要的一步.因为以后再研究这类空间,只需考察一个空间的下空间就可以了. П.С. Александров[10,19]定理说:

任予超穷基数 τ,空间 F^{τ} 是重$\leqslant\tau$ 的 Колмогоров 空间类的万有空间.

A.H. Тихонов[4]又另外做出 T_0 类的万有空间.

11. Riesz 空间范围比较狭一点,它的特征是下列公理:

任两不同一点集的每个集近傍分离于另一个.

换句话说,x,y 是空间两不同点,于是点 x 有近傍不含点 y,或者又可以说:所有一点集是闭. Reisz 空间类用符号 T_1 表示.它包含在 T_0 类,但不相等.譬如,上述 F 属于 T_0 而不属于 T_1. T_1 类具遗传性及乘性.重$\leqslant\tau$ 的 Riesz 空间类的万有空间是否存在的问题至今尚未解决.

12. 再狭些的类叫做豪斯道夫空间.它的特征是下列公理:

任两不同一点集互相近傍分离.

豪斯道夫空间类用符号 T_2 表示.它属于 T_1,但不相等. T_2 类具遗传性及乘性.

13. 关于 T_2 类,提出了很丰富及有兴趣的问题,苏联拓扑学者对于这些问题的探讨贡献很大.我们引进下列术语,以便叙述其中一个有关的问题.

设 R 是某一空间类,若空间 X 属于 R,并且在所有含 X 为下空间的 R 类的空间里闭的时候,我们说空间 X 绝对闭于 R.

我们想求绝对闭于 T_2——或叫做 H 闭——的空间的内在特征(对于 T_0 及 T_1 的相仿 7 问题,很简单,无须提及).

H 闭性问题由 П.С. Александров 及 П.С. Урысон 的下列定理所解决:

豪斯道夫空间 X,H 闭的充要条件是从 X 的任一个遮里,可以选出有穷个集(遮的元),其和稠于 X.

这里及以后常用空间 X 的遮,它的意思是:空间 X 的闭集系,其和等于 X.

14. 空间 X 若是空间 Y 的下空间,而稠于 Y 的时候,我们说 Y 是 X 的推广.

П.С. Александров 及 П.С. Урысон[3]提出了有关 H 闭的问题:任一豪斯道夫空间是否有 H 闭的推广? 头一个接近这个问题的是 A.H. Тихонов[2],他证明了:所有豪斯道夫空间是某些等重 H 闭空间的下空间.决定性的并且是肯定的解答是由美国数学家 Stone 完成的. Stone 的定理说:所有的豪斯道夫空间恒

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

有 H 闭推广. 他的定理后来经 А. Д. Александров[3]、С. В. Фомин[1] 及 Н. А. Шанин[1]用各种不同的方法重新证明. С. В. Фомин 并且同时证明了所有豪斯道夫空间有等重 H 闭的推广. Stone 定理的各种证明中, 概念最简明的是 А. Д. Александров的证明.

15. 任予基数 τ, 最小的基数 σ 能满足不等式 $\tau \leqslant 2^{\sigma}$ 的叫做 τ 的对基数. М. Я. Перельман①证明了下列有意思的定理:

无穷豪斯道夫空间系所作成的积空间的拟重等于因子空间的拟重及这些因子的个数的对基 数的上限, 但是假设每个因子空间都要多于一点.

16. H 闭这个概念与"紧"这个概念有密切的关系. 紧的意思是: 从所有的遮里恒可以选出有穷个集元, 仍能构成空间的遮.

从波莱尔 - 勒贝格定理, 我们知道所有 n 维数空间的囿、闭下集看做下空间的时候是紧. 在分析上囿、闭集之如何要紧是周知的. 所以很自然的, 我们可以预料在拓扑学以及它的应用里, 紧空间想必也有相仿的重要性. 这个忖度很对, 紧空间曾经是 П. С. Александров 及 П. С. Урысон[2,3,4]所建立的广大理论的对象. 这理论的中心思想——紧——或对于拓扑学全体, 或对于许多它的应用都显示着本质性的功效. 我们打算在下面讨论它的一些性质.

17. П. С. Александров 及 П. С. Урысон[2,3]曾建立上述紧的定义与其他两个定义之间的同值.②

其中一个是这样的: 空间里所有不空、闭集的逆减良序列都有不空的交的时候, 称空间为紧. 由这个定义所表示的紧的解释, 我们知道是那有名的嘉当的关于交的定理的一般化.

另外一个紧的定义则与"全聚点"有关. 设 A 是空间 X 的下集, x 是 X 的一点. 若 A 与 x 的任一近傍的交的势能与 A 的势相等的时候, 我们称 x 做 A 在 X 的全聚点.

相关的紧的定义这样说: 空间的任一无穷下集一定有全聚点的时候叫做紧空间. 这个解释, 是波尔查诺 - 魏尔斯特拉斯定理的一般化.

波尔查诺(Bolzano, Bernard, 1781—1848), 捷克哲学家、数学家、逻辑学家、宗教家、伦理学家.

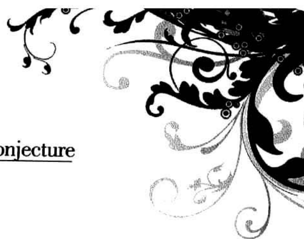
18. 紧空间构成乘性类. А. Н. Тихонов[2]证明的这个结果决不是句空话③. 从这个定理, 特别可以得出, F^{τ}, D^{τ}, I^{τ} 都是紧空间.

19. 从紧空间的定义马上得出, 对于连续变换, 紧性是被保持的. 这就是 Н. Ь. Веденисов的定理: 若能有自紧空间 X 到空间 Y 上的连续变换的时候, Y 也是紧空间.

比较一下紧空间的定义, 同 П. С. Александров 及 П. С. Урысон 关于 H 闭

① Мощности в топологии. Л., Рукописв(拓扑学里的势, 稿本), 1941: 1-361.
② 参看 П. С. Александров[19].
③ А. Н. Тихонов 只考虑了线段系的积, 不过他的论断可以逐字适用于一般情形: 紧空间系的积.





定理间的定理,可以得出:所有豪斯道夫紧空间都是 H 闭.

紧空间类不具遗传性.但是不难看出:紧空间的所有下空间,看做下集,若闭于 X 的时候,一定是紧空间.

把所有这些事实联系起来,就得到 Н.Ь.Веденисов 的定理:

设有连续变换自紧空间 X 到豪斯道夫空间 Y 内,这时所有空间 X 里闭集的象闭于 Y.

从这里直接得出来:自紧空间到豪斯道夫空间内的所有(1,1)连续变换都是拓扑变换.

我们注意一个很有意思的事:这最后的结果同它在局部的翻译,也就是关于连续于一点的变换一起构成 А.Н.Тихонов 定理[6]的拓扑学方面的基础.他证明了某些分析上反值问题的稳定性,特别是关于位论的反值问题.

确有 H 闭而不紧的空间,П.С.Александров 同 П.С.Урысон[3]曾作了这种空间的例.

紧豪斯道夫空间我们叫做紧体.

20.与紧空间同时,П.С.Александров[2]又引进了"局部紧"空间,它的定义是这样的:

x 是空间 X 的点,假若 x 有在 X 的近傍,并且其闭包紧的时候,我们说空间 X 紧于点 x.若 X 紧于它的每个点的时候,我们说 X 局部紧.

所有紧空间,显然是局部紧,但另一方面,有局部紧,而不紧的空间,例如,n 维数空间就是.

П.С.Александров[2]的基本结果,可以使我们将考虑任一个局部紧豪斯道夫空间归为考虑某些紧体.这紧体是由添一个新点于原来的局部紧空间得到的.这手续的特殊情形,例如添一个"无穷远"点于局部紧而不紧的复数平面,于是得出紧的 Gauss 球.

要想叙述 П.С.Александров 的定理,先得引进下列术语.含空间 X 为下空间,与 X 只差一点的紧体,我们叫做空间 X 的一点紧化,П.С.Александров[2]定理可以叙述如下:

空间 X 的一点紧化存在的充要条件是 X 是豪斯道夫局部紧空间.这时 X 的一点紧化实在是唯一的.

由这里可以得出局部紧豪斯道夫空间的特征.可以这样说:它是紧体的开下集而被看做空间的.

21.П.С.Александров[1]引进了一个很重要的概念"势敛".定义是:空间 X,其无穷下集 A,点 x,任取 x 的近傍 U,若 $A \setminus U$ 的势比 A 的势小的时候①,我们

① "$A \setminus B$"表示集 A,B 的差,就是说不属于 B 的 A 的元全体,"$A \cap B$"表示 A,B 的交,"$A \cup B$"表示 A,B 的和.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

说 X 的下集 A 势敛于点 x. 任予点 x，要想有下集 A 势敛于这点，当然这点 x 必须不是孤点. X 若是豪斯道夫空间，并且在 x 紧的时候，这个条件又是充分的. 也就是说，我们有下列的并不是自明的 П.С.Александров[1]定理：

豪斯道夫空间 X 紧于它的非孤点 x，于是 X 有无穷下集 A 势敛于点 x.

22. 设 U 是空间 X 的点 x 的近傍所成系. U 满足下列条件：点 x 的任一近傍一定含某些属于 U 的近傍的时候，我们说 U 是 X 在点 x 的基.

例如点 x 在 X 的所有的近傍所成系当然就是 X 在点 x 的基. X 在点 x 的基的势的最小的，叫做 X 在点 x 的重. X 在点 x 的重，用符号 $p_x(X)$ 表示.

空间 X 的点 x 的近傍系 U 的所有集元的交，只是一点 x 的时候，我们说 U 是 X 在 x 的拟基.

X 若是 T_1 类的空间的话，于是这空间的所有的点都有拟基：例如点 x 的所有的近傍，就是 X 在 x 的拟基. 在这种情形，X 在 x 的拟基就是 X 在 x 的基. X 在点 x 的拟基的势的最小的，叫做 X 在点 x 的拟重. X 在 x 的拟重，用符号 $q_x(X)$ 表示.

显然，在 T_1 类空间的时候，$q_x(X) \leqslant p_x(X)$.

П.С.Александров[1]证明了下列定理：

X 是豪斯道夫空间，并且紧于点 x 的时候，$q_x(X) \leqslant p_x(X)$.

22. 对于空间 X 的点 x，若能有下集 A 势敛于 x，那么在这些下集的势里一定有最小数. 我们把它叫做 X 在点 x 的势重. 空间 X 在它所有孤点的势重，我们定义为 1. X 在 x 的势重，用符号 $c_x(X)$ 表示.

由上述 П.С.Александров 定理，所有豪斯道夫空间，紧于它的某一点 x 的时候，一定有确定的势重. 在这种情形，П.С.Александров 证明了 $c_x(X) \leqslant p_x(X)$.

24. 关于紧空间的许多定理，可以推广到更广泛的“列紧”空间，它的定义是这样的：若能自空间 X 的可数遮选出有穷个集元，仍足以遮 X 的时候，定义 X 是列紧. 同紧空间一样，这时也可以有不相同而同值的定义. 其中一个是这样：空间的任一递减不空闭集列，必有不空的交的时候，这空间是列紧.

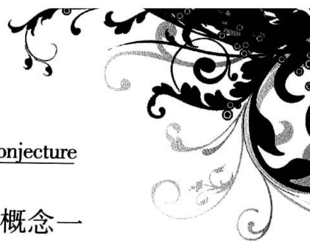
另外一个定义，则与“聚点”有关. 设 A 是 X 的下集，x 是空间的一点. 若 A 与 x 的任一近傍的交都是无穷的时候①，我们说 x 是 A 在 X 的聚点.

用这个概念定义列紧就是：空间的任一无穷点集必有聚点的时候，空间是列紧.

所有紧空间，显然，列紧. 另一方面，很容易看出：所有列紧空间，它的重不

① X 是 T_1 类空间的时候，代替原条件以要求 A 与 x 的任一近傍的交至少含一点不等于 x 也可以. 在一般情况，如此改变定义，则得出完全新的概念.





超过可数的时候必是紧.因此,空间的重不超过可数的时候,紧与列紧两概念一致.

但是一般这两概念是不一致的:因为确有列紧而不紧的空间(П.С. Александров 及 П.С.Урысон[3]).

25.与空间紧于某点相仿的概念是“列紧于某点”.意思说,X 的点 x 有一适当近傍其闭包列紧.若空间 X 列紧于它所有的点,我们说 X“局部列紧”.

与关于空间的一点紧化的定理相仿的是局部列紧空间的一点列紧化的定理(П.С.Александров[2]):

空间 X 有一点列紧化的充要条件是 X 是豪斯道夫局部列紧空间.

这里所提到的空间 X 的一点列紧化是指一个列紧豪斯道夫空间,含 X 为下集,并且与 X 只差一点.定理的条件,不能保证一点列紧化的唯一性.

26.要在一个空间施行分析的话,必须在这空间有“足够”的实值连续函才好.在 T_2 类空间,有时或许“稍少”.只取一个值的,空间的实函,显然是连续的.关于所有的空间,很自然地要产生这个问题:就是是否在空间里,除这些无谓的以外还有连续函.我们晓得有无穷多个 T_2 类的空间,在它的实连续函只能是常数.头一个,这种例是 П.С.Урысон[4]作的.这个例的空间是可数.1937 年捷克数学家 Pospisil 推广 Урысон 的做法,证明了任予无穷基数 τ,一定有势是 τ 的豪斯道夫空间,在它定义的实连续函必是常数.在这些例里,不是常数的连续函之所以不存在是根据做出来的空间的下列性质:任两不空开集的闭包必相交.具这性质的空间显然通.所以 П.С.Урысон 的例同时是一个可数、通、豪斯道夫空间的例.

27.П.С.Урысон 的例告诉我们,T_2 类空间,对于施行“分析”太宽泛.因此就得在考虑的空间上加更重的分离条件,很自然地缩小研究的范围.

其中的一个就是条件“常”:

任一点 x 及不含此点的开集 A,$\{x\}$ 及 A 互相近傍分离.

T_1 类的空间,而能满足这个条件的时候,叫做常空间.常空间类用符号 T_3 表示.

很容易看出来,所有的常空间是豪斯道夫空间.另一方面,确有不常的豪斯道夫空间.例如上面提到的 Урысон 作的空间就是.很容易看出,T_3 类具遗传性及乘性.

28.П.С.Александров 及 П.С.Урысон[2,3]证明了下列紧体的特征:

空间是紧体的充要条件是空间常而且 H 闭.

这样,在 T_3 类,紧体与 H 闭两概念一致.对于这一类,发生绝对闭的问题——就是说找绝对闭于 T_3 类的空间的内在特征.这问题经 Н.М. Вайнберг[1]所解决.他是彼得格勒的数学家,在伟大卫国战争里壮烈牺牲了.

要想叙述他的结果，我们引进下列术语．空间的遮的每一个元的闭包必定含于遮的另外一个元的时候，我们说这是空间的 r 遮．Н.М.Вайнберг[1]的定理是这样的：

常空间绝对关于 T_3 的充要条件是从它的所有 r 遮里，可以选一有穷遮．

29．П.С.Александров 及 П.С.Урысон[3]会证明了下列定理：

所有列紧豪斯道夫空间 X 满足 $p_x(X) \leqslant \aleph_0$ 于空间的常点 x．

列紧常空间的全集的势不少于 $2^{\aleph_0}$．

第二个定理特别与紧体有关．

П.С.Александров[1]会考虑关于紧体里全集的存在问题．我们只从他得到的结果列举一个．

任予基数 τ 小于 X 的基数，基在点 x 的所有近傍里，有点 y 而满足 $p_y(X) > \tau$的时候，我们设点 x 是空间 X 的特点．于是成立下列定理：

所有不含全集的紧体一定有特点，孤于特点的全体．

30．用拓扑学去接近分析的时候，T_3 类仍觉过于宽泛．为了这个，我们要求更强的分离条件，例如下列的条件“法”．

任两个不相交的闭集互相近傍分离．

T_1 类空间，更满足这个条件的，我们叫做法空间．法空间类，用符号 T_4 表示．

T_4 类显然含于 T_3 类．反包含并不成立．有如 П．С．Александров 及 П.С.Урысон[3]所例示的．

法空间的理论是由 П.С.Урысон 所建立的．这理论的最重要结果就是下列连续函推广的重要定理（П.С.Урысон[3]）．

设 A 是法空间 X 的闭集．所有 A 上囿、连续实函可以推广成 X 的连续函．

这个定理是有名的豪斯道夫的关于计量空间（见 35 段）的定理的一般化．假设我们停留在 T_1 类范围之内的话，П.С.Урысон 的定理并且更进一步不允许我们推广所考虑的空间．因为很容易证下列 П.С.Урысон 的反定理：

空间 X 是 T_1 类，并满足下列条件：任取空间 X 的闭集 A，所有 A 的实、囿、连续函都能推广成 X 的连续函，于是 X 是法空间．

虽然如 А.А.Марков[3]所指出，空间之属于 T_1 类，对于适用连续函推广的定理决不是什么不可缺少的条件．因为 П.С.Урысон 所给的这个定理的证明对任何空间都成立，只要它的任两不相交的闭集能互相近傍分离就行．最后 А.Д.Александров证明可以把 П.С.Урысон 定理一般化，把它推广到前述满足 $E_2{}^*$，不要 E_2 的空间．在这样的 А.Д.Александров 空间里，只要任两闭集不相交，则互相近傍分离，就可以应用 Урысон 定理推广连续函（А.Д.Александров[1]）．





Н.Б.Веденисов[6]在另一方向推广 П.С.Урысон 定理,证明了被推广的函的囿这个要件可以不要.

П.С.Урысон 定理告诉我们,法空间里连续函够多.特别由这定理可以得出,在法空间,只要它有一个以上的点,一定有不是常数的连续函.

П.С. Урысон 定理允许我们在法空间作"分析"及其推广.特别 А.А.Марков[3]在法空间构成了一般测度及积分的拓扑理论.这理论经过 А.Д.Александров[1,2,4]的工作更得到重要的发展.П.С.Урысон[4]从他的定理,得到了重要的系,这系是关于法空间的通集的势的:

法空间的通下集,只要它含一个以上的点,它的势$\geqslant 2^{\aleph_0}$.

对于常空集的通下集,这定理未必成立.П.С.Урысон[4]只证明了:所有常空间的通下集的势,只要它含一个以上的点,就不是可数.

31.在拓扑学以及它的应用里,法空间既是如此的重要,找些另外的条件以断定其他的空间能是法空间的定理,因此也有重要的意义.А.Н.Тихонов[1]的定理就是这样定理中的一个:

所有可数重的常空间是法空间.

另外这样的定理是 П.С.Александров 及 П.С.Урысон[2,3]的:

所有紧体是法空间.

32.T_4 类既不具遗传性也不具乘性.

А.Н.Тихонов[2]作了一个简单的例,说明由紧体去掉一点就得到含于紧体的下空间,而不是法空间.

法国数学家 Dieudonné 作了一个例,说明由两个法空间所得积,可以不是法空间.

33.上面已经注意过,任予超穷的 τ,I^τ 的重是τ.对任一 τ,I^τ 是紧,并且因为 T_2 类具乘性,所以属于 T_2 类.空间 I^τ 既是紧因之是法空间,特别 $I^{\aleph_0}$是可数重的法空间.П.С.Урысон[2]证明了下列很有意义的结果:

空间 $I^{\aleph_0}$是可数重法空间的万有空间.

根据上面所提到的 А.Н.Тихонов 定理,我们可以把这里的"法"字改成"常"字,以便说明所考虑空间类的遗传性.

为了对任意重的空间,得到相仿的结果,А.Н.Тихонов[2]引进了又一个重要的空间类——"全常"——空间.全常空间定义与常空间定义相仿,不同于后者只是代近傍分离以泛函分离.

空间 X 属于 T_1 类,并满足下列条件的时候,叫做全常:行予空间 X 一点 x 及闭集 A 不含 x,集$\{x\}$与集 A 必定泛函分离.全常空间类用符号 T_ρ 表示.

很容易看出来,T_ρ 类具遗传性.由定义,马上看出 T_ρ 类含于 T_3 类.从连续函推广的定理得出 $T_4 \subset T_\rho$.另一方面,А.Н.Тихонов[2]作了一个常而不是全常

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

空间的例.上面证明 T_4 空间不具遗传性的例,同时是全常而不是法空间的例.所以 T_ρ 类位居 T_3,T_4 两类之间,比前者窄,比后者广.重$\leqslant \tau$ 的全常空间类是具遗传性的,Γ^τ 就属于这类.原因是 Γ^τ 的重$\leqslant \tau$,并且是法空间,因之是全常.A.H.Тихонов[2]证明了下列有意义的定理:

任予超穷基数 τ,Γ^τ 是重$\leqslant \tau$ 的全常空间类的万有空间.

因此全常空间的研讨,都归成 Тихонов 方体的研究.

由 A.H.Тихонов 的定理直接得出一个重要的系,关于看全常空间为紧体的下空间,或看做法空间的下空间的特征的:

空间是紧体的下空间的充要条件是全常.

空间是法空间的下空间的充要条件是全常.

所有这些结果充分证明了全常空间在拓扑学的意义.目下在数学的很多邻近分支里,发现这一类空间具有很重要的意义,例如在拓扑群的理论①.原因是Л.С.Понтрягин 及 A.A.Марков 得到了下列定理:

空间是拓扑群的闭下空间的充要条件是全常.

A.A.Марков[4]提示了许多拓扑群构成的一般方法,这时全常空间都是非常重要的.

34.从上述 A.A.Тихонов 的结果,得出所有全常空间都有紧,豪斯道夫推广.与这些结果相结合,捷克数学家切赫,更进一步发展这思想,证明全常空间 X 的所有豪斯道夫,紧推广里,有按下列意义最大的一个:它可以连续变换到所有 X 的豪斯道夫,紧推广上,并且 X 本身的点仍变换到本身上去.这切赫推广实际是唯一的,以后用符号 βX 表示它.切赫更证明 βX 有下列性质:所有自空间 X 到任意的紧体 Y 内的连续变换可以推广为自 βX 至 Y 内的连续变换.很容易看出来,上述 βX 的最大性可以由此得出.

1938 年美国数学家 Wallman 提示了任意 T_1 类空间紧推广的另一方式的建设.这个及切赫的构成,读者可以参考 П.С.Александров[19].空间 X 的 Wallman 推广用符号 ωX 表示.

实质上新,并且由某种观点看更简单的构成要算 П.С.Александров[12].同时他解释了全常空间 X 的推广 βX 及 ωX 之间的关系.βX 可以由 ωX 按一定的方式得出来.

1940 年 A.Д.Александров[1]将 Wallman 的紧推广一般化为空间的任意推广,并且证这构成的结果——空间 X 的紧推广 ωX——永远满足下列性质:

自空间 X 至任意紧体 Y 内的所有连续变换,可以推广为自 ωX 至 Y②内的连续变换.

---

① 关于拓扑学的定义,读者请参看《苏联数学三十年》内拓扑学,Lie 群项.
② 此处只引了 A.Д.Александров 的结果中与普通拓扑空间概念一致的一部分.





更继续研究紧推广的理论,А.Д.Александров[1]证明了一连串其他极有兴趣的推广.其中有些与"至法"空间这概念有关.这概念是 Н.Б.Веденисов 引进的,其定义如下:空间的所有闭集是由某些在 X 的实连续而在该处趋于零的点所作成的时候,叫做至法空间.Н.Б.Веденисов[1]证明了下列定理:

空间是至法的充要条件是它是法空间,并满足所有它的闭集是不超过可数多个开集的交.

至法空间这概念及 Н.Б.Веденисов 定理经 А.Д.Александров[1],撤销了属于 T_1 类的限制,并把结果推广到他的一般空间(见第一段).至法空间这概念,在 А.Д.Александров 的拓扑测度论里有极重要的意义.

Н.А.Шанин[1,2,3]更进一步很有成就地研究了有关推广的理论.

35.苏联拓扑学派伟大成就之中的又一个,是"计量化问题"的解决.这问题与"计量空间"这概念有关,其定义如下:

设 X 是任一集,ρ 是 X 的两元的实函,满足下列两条件:

D_1　$\rho(x,y)=0$ 只限于 $x=y$.

D_2　任予 $x,y,z\in X$,$\rho(x,y)+\rho(y,z)\geqslant\rho(z,x)$.

我们说 ρ 是 X 的计量,所谓计量空间就是同时考虑这个集及在这集上定义的计量.计量空间的元叫做它的点.数 $\rho(x,y)$——ρ 是计量空间 X 的计量,x,y 是 X 的点——叫做 x,y 在 X 的距离.

从条件 D_1,D_2 得出,任两点之间的距离不负,并且 $\rho(x,y)=\rho(y,x)$,并且对任三点 x,y,z,$\rho(x,y)+\rho(y,z)\geqslant\rho(x,z)$.

设 x 是计量空间 X 的点,ε 为任一正数.与 x 的距离小于 ε 的所有的点 y,叫做 x 在 X 的 ε 近傍.

设 A 是计量空间 X 的下集.任取 A 的点 x,于是有 $\varepsilon>0$,x 的 ε 近傍全含于 A 的时候,我们说 A 是开.

计量空间的开集满足 E_1,E_2,E_3.因此计量空间确定一个拓扑空间.拓扑空间从计量空间出发,由这种方式决定的了的,我们叫做可计量空间.所有拓扑空间都可计量是决不能成立的.但可以知道:所有计量空间是法空间.

计量化问题就是求可计量空间的内在特征,也就是用纯拓扑的语言来表达可计量空间的特征.这问题由 П.С.Александров 及 П.С.Урысон[1]所解决.为了叙述他们所得的结果,我们引进下列术语.

α,β 是空间 X 的两个遮,任取 α 的两元 U,V,而 $V\cap V\neq\Lambda$,于是有 β 的元 W 满足 $U\cup V\subset W$ 的时候,我们说 α 是嵌于 β.空间 X 的遮列 $\alpha_1,\alpha_2,\cdots$,若满足:任取点 x,及集 $U_1,U_2,\cdots$,涉 $x\in U_i\in\alpha_i(i=1,2,3,\cdots)$,则集系 $\{U_i\}$ 是 X 在 x 的基的时候,我们说遮列 $\{\alpha_i\}$ 是在 X 的全锁.在 X 的全锁 $\{\alpha_i\}$ 若每个 $\alpha_i\alpha_{i+1}$ 嵌于 α_i 的时候,我们说 $\{\alpha_i\}$ 常.

П.С.Александров 及 П.С.Урысон 所发现的计量化问题的解答是这样的:

空间可以计量化的充要条件是空间有常全锁.

与这个计量化一般定理不相关,另有三个关于特殊不同空间类可以计量化的定理.

其中头一个定理是 П.С.Урысон[5]所给的可数重空间计量化的充要条件:

空间的重不超过可数,其计量化的充要条件是它是常空间.

这定理很容易可以由上述 П.С.Урысон 关于空间 $I^{\aleph_0}$的万有性得出来.因为注意这空间拓扑同值于希尔伯特空间的方体就可以.方体按定义是:$0\leqslant x_i\leqslant \frac{1}{i}(i=1,2,\cdots)$,此处 x_i 是希尔伯特空间的点的坐标(А.Н.Тихонов[2]).

第二个计量化定理,也是 П.С.Урысон[1]所证明的,是关于列紧空间的.

列紧空间可以计量化的充要条件是它是豪斯道夫空间,而重不超过可数.

必要部分,早由豪斯道夫所证明,充分部分,很容易由第一计量化定理得出.因为只需注意所有列紧,而重不超过可数的豪斯道夫空间是紧体,因之,是常空间.

列紧计量空间叫做列紧体.

П.С.Александров 所证明的第三个计量化定理是关于局部列紧空间的:

局部列紧空间可以计量化的充要条件是它是豪斯道夫空间,并可表示成互斥开集之和,每个开集的重不超过可数.

只是必要部分的证明相当难.

36.拓扑学重要问题中的一个是研究"什么空间可能?"的问题.当考察这个拓扑学问题的时候,很自然地,应将拓扑同值的空间抽象地看做一个.

很明显的,这最一般的问题很不容易处理:因为空间太多,太繁杂,不便于研讨它的全体.因此适当地限制空间类,不损伤原问题的精神,以便仔细地研究,应能很有成果.特别第一步应当考察我们所考虑的类里,拓扑不同值的空间表现的计算.也就是说弄清这表现全体的势.М.Я.Перельман①作了很多这类的计算.

下面所引的四个 М.Я.Перельман 的定理里的"τ"表示任意的超穷基数.

恰有 2^{τ} 个重是 τ 的拓扑不同值的紧体.

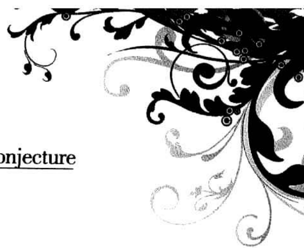
恰有 $2^{\tau\aleph_0}$个重是 τ 的拓扑不同值的可计量空间.

恰有 $2^{2^{\tau}}$个重是 τ 的拓扑不同值的 T_1 类空间.拓扑不同值的、拟重是 τ 的全常空间亦同.

恰有 $2^{2^{2^{\tau}}}$ 个拟重是 τ 的拓扑不同值的常空间.拓扑不同值的、拟重是 τ 的全

① Мощности в топологии.Л.,Рукописв(拓扑学里的势,稿本),1941:1-361.





常空间亦同.

37. П.С.Александров[5]证明了下列重要定理:

嘉当不连续集可以连续变换于任一列紧体上.

与列紧体是紧体的一部分一样,嘉当不连续集 $D^{\aleph_0}$是一般嘉当不连续集 D^{τ} 的特例.因此可以预想,任一紧体可表示成一般嘉当不连续集的连续象.这个预想为波兰数学家 Spilrjagn 所否定.紧体而可以表示成空间 D^{τ} 的连续象的,П.С.Александров 把它叫做二进紧体.Н.А.Шанин[4,5,6,7]对它的理论作了很深刻的研究.许多以往对于列紧体成立的定理,都可以推广到二进紧体.为了叙述 Шанин 的主要结果,我们引进下列术语.集系中任两元 A,B 之间必有 $A\subset B$ 或 $B\subset A$ 成立的时候,我们称之为序系.Шанин 证明了下列定理:

二进紧体 X 的稠于 X 的开集所成不空序系的所有的集的交仍稠于 X.

由此得出,不空二进紧体不能表示成疏于 X 的下集所成序系的所有集之和.

Н.А.Шанин 作了一个反例,进一步证明,上列结果中二进性是去不得的:确有紧体 X,是疏于 X 的闭下集所成序系的所有的集之和.

Н.А.Шанин 另外得到关于“序”二进紧体的有趣结果.按下述方法作成空间,叫做序空间.设 X 是任一序集,集 X 的隔间是合乎下列四种情形之一的 X 的下集:

(1)所有位于 X 的两固定元之间的元全体.

(2)所有位于 X 的某固定元之前的元全体.

(3)所有位于 X 的某固定元之后的元全体.

(4)整个集 X.

由隔间所作的和叫做开集,于是 X 成法空间.Н.А.Шанин 证明了下列定理:

所有序二进紧体拓扑同值于数直线的囿、闭下集.

如有名的例所指出,此处二进这条件仍旧是去不得的.

与列紧体是一个固定的空间 $D^{\aleph_0}$的连续象相仿,是否一般紧体也可以表示为特别空间的连续象,П.С.Александров[10,19]会证明下列关于这方面的定理:

所有重$\leqslant\tau$ 的势体是空间 D^{τ} 的某些下集的连续象.

同时证明了关于所有 T_0 类的下列结果:

所有重$\leqslant\tau$ 的 T_0 类空间是空间 D^{τ} 的某些下集的(1,1)连续象.

38. 上面所引 Н.А.Шанин 的结果是根据他的“口径理论”,这理论有它独立的兴趣.这理论的中心思想——拓扑空间的“口径”——的定义是:基数 $\tau>1$,所有空间的开集系,势是 τ,必定含同势的下系,其中开集之交不空的时候,我们称基数 τ 为空间 X 的口径.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

Н.А.Шанин 研究了积空间的口径与因子空间口径之间的影响.得到的这些结果,使他能够证出上述关于二进紧体的诸定理.

39.1928 年 П.С.Александров[7] 发表了第一篇他的“谱论”的文章.这理论的基本趋势是从其他简单明了的对象出发,应用些极限算法做出来的极其一般的空间.换近,例如,在 П.С.Александров[19] 的论文里的处理方式,出发的材料是陨空间.所谓陨空间是 T_0 类空间,其所有开集之交仍是开集.很明显的,所有有穷 T_0 类空间是陨.

在陨空间,所有集系的集之和的闭包等于集系的集的闭包之和.但每个点集是一点集之和,所以在陨空间一点集的闭包若已知,则闭包这运算完全确定.

陨空间 X 中的两点 x,y 间的关系

$$x\in \overline{\{y\}}\setminus\{y\} \qquad ①$$

在 X 决定一个反对称,可递的关系,也就是一个半序①.

不难看出,上列关系决定一个陨空间与半序集之间的(1,1)对应:所有陨空间决定一个半序集,这半序集由空间的点所构成,“x 在 y 前”是用条件①定义的;所有半序集按这个意义决定唯一的一个陨空间.为了由半序集 X,得到唯一的陨空间,必须将满足下列条件的 X 的下集看做开:集 X 的元 x 若属于 X 的某集,而所有 X 的元在 x 后的也属于该集.陨空间的理论,这样一来,就同值于半序集的理论.

以下式子“$x<y(X)$”表示,在半序集 X 里,x 在 y 前;式子“$x\leqslant y(X)$”表示 x,y 是半序集 X 的元,$x<y(X)$或 $x=y$.

设 f 是自半序集 X 至半序集 Y 的变换,若 $x\leqslant y(X)$,必定 $f(x)\leqslant f(y)(Y)$ 的时候,我们说 f 保序.

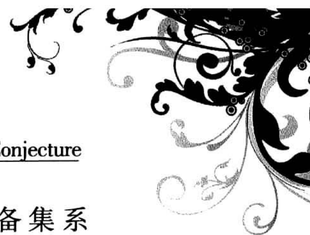
很容易看出来,自陨空间 X 至陨空间 Y 内的变换连续的充要条件是定义在 X 的半序经过这变换保序于定义在 Y 的半序.

任一半序集 X 对应一个反半序集 X^*,X^* 的元与 X 完全相同,只是 $x<y(X^*)$时,必须而只需 $y<x(X)$.转移到反半序这个运算相当于作予陨空间的匹空间.所谓匹空间的意思是,点与原来的一样,但匹空间的开集正是原来的闭集.

任意的集系可以看做半序集,注意其包含就具有半序的性能,就是说 A,B 是集系的元,若 $A\subset B$ 而 $A\neq B$ 的时候,我们定义 A 在 B 前.既如此,再根据半序集与陨空间的同值,所有集系都可以看做是陨空间.另一方面,不难看出,所有半序集同构于某些自然的半序集.由此得出,陨空间拓扑同值于视为陨空间的集系.

① $\bar{A}$ 表示集 A 在所考虑空间的闭包.





若一个集系，其每一个元的下集也都是系中一元，则称为备集系(П.С.Александров[11]).有穷备集系称为抽象单纯复合形，简称为 A – 复合形.

显然，抽象单纯复合形中的元是有穷集，由于 A – 复合形与欧几里得单纯复合形间的关联，П.С.Александров 称为穷集为间架，称具有 $n+1$ 个元素的集为 n 维间架($n=0,1,2,\cdots$).这样，每一个 A – 复合形就看成了间架所组成的有穷集.若 C 是一个 A – 复合形，包含有 n 维间架而不包含任何 m – 维($m>n$)间架，则称 C 为 n 维 A – 复合形.

A – 复合形的每一个下集是一个集系，且可以看做一个有穷陨空间，即一个 T_0 类中的有穷空间.反之，如同 П.С.Александров[11]，每一个有穷陨空间必同胚于某一 A – 复合形的下集.这样一来，对于有穷陨空间的研究就化为对于 A – 复合形的研究，越加显出了这种空间理论的组合本质.

现在我们转到对于 П.С.Александров[7,9,19] 及 А.Г.Курош[1,2] 所创立的“谱论”的简明叙述，这个理论给予了由有限陨空间出发来构成紧体的可能性，并且开辟了用组合代数方法来研究紧体的一条要道.

40.半序集 Ξ 假如满足下列的条件，则称为有向：对于集 Ξ 中的任意二元 α 与 β，必存在 Ξ 中一元既随在 α 之后，又随在 β 之后①.

设对于有向集 Ξ 中的任意一个元 α 都对应了一个空间 X_α 并且对于 Ξ 中的每一对元“α,β”适合 $\alpha<\beta(\Xi)$ 的，都对应了一个由空间 X_β 到空间 X_α 的连续变换 π_α^β，满足下列关系

$$\pi_\alpha^\gamma=\pi_\alpha^\beta\pi_\beta^\gamma$$

只要是 $\alpha<\beta<\gamma(\Xi)$.空间系 X_α 与变换系 π_α^β 合并起来考虑，称为一个谱.设谱由空间 $X_\alpha(\alpha\in\Xi)$，及变换 $\pi_\alpha^\beta(\alpha<\beta(\Xi))$ 组成.看空间系 $X_\alpha(\alpha\in\Xi)$ 的积 X.X 中的点是所有可能的系 $\langle x_\alpha\rangle_{\alpha\in\Xi}$，其中 $x_\alpha\in X_\alpha(\alpha\in\Xi)$.令 Q 表示满足下列条件的系的全体：$\pi_\alpha^\beta x_\beta=x_\alpha$，只要是 $\alpha<\beta(\Xi)$.作为空间 X 的下集，Q 也是一个空间.空间 Q 称为所考虑的谱的备限.由不空紧体所产生的谱，其备限是一个不空紧体(П.С.Александров[19]).

从有限陨空间来构成紧体的过程中，П.С.Александров 所引进谱的低限的观念担任了一个重要角色，它可以定义为谱的备限中那种点 $\{x\}$ 的全体，满足条件 $\overline{\{x\}}=\{x\}$，就是说，一点集 $\{x\}$ 是闭集.此处，闭包是在备限中来取的，所取的备限的下集看做一个空间②.

显然，每一个谱的低限是 T_1 类中的空间，П.С.Александров 陈述了下面的充要条件使得谱的低限是 T_2 类中的空间：对于谱的低限中任意两个不同的点

① 我们宁可以“有向集”这个术语来代替具有相同意义而为 П.С.Александров 所采用的术语“无界集”.

② П.С.Александров[19] 中，低限的定义是用另外的术语来陈述的.

$<x_\alpha>_{\alpha\in\Xi}$与$<y_\alpha>_{\alpha\in\Xi}$,必存在$\alpha$使得在$X_\alpha$中$x_\alpha$与$y_\beta$是近傍分离的.满足这个条件而且是由有穷陨空间所产生的谱,П.С. Александров 称之为豪斯道夫谱①.

谱论的主要结果是下面的 Александров-Курош 定理:

每一个豪斯道夫谱的低限必为一紧体,并且每一个紧体必为某一豪斯道夫谱的低限(П.С. Александров[19]).

41. 谱论使得有可能实现在逻辑上自然地造出"n – 维元体"——多面体拓扑学中的结构单位.所谓 n – 维元体——简称 n – 元体——是指所有与 n – 维方体 $I^n(n=0,1,2,\cdots)$同胚的空间,I^0是定义为只有一点的空间.

П.С. Александров 在构成以 n – 元体为低限的豪斯道夫谱的过程中,用到了半序集"重心派生"的观念,其定义如后.半序集 X 所有的不空下序集全体在自然的意义之下组成了一个半序集 X_1,称为 X 的重心派生.重心派生永为备系.有限半序集 X 的重心派生显然也是有穷集,因此是 A – 复合形.在这种情形,如同在任何有穷半序集一样,重心派生中的每一元 Y 中必含有一个末元(在集 Y 的所有其他元之后).集 Y 的这个末元称为 Y 的负荷元.有穷半序集 X 的重心派生 X_1 的诸负荷元都是 X 本身的元.将重心派生中的每一个元对应于它的负荷元,我们就得到自半序集 X_1 到半序集 X 的一个变换,容易看出这个变换是保序的.我们称它为重心派生的自然变换.

利用半序集与陨空间之间的对应,可以这些观念推及有穷陨空间.可以定义有穷陨空间 X 的重心派生以及将它变换到 X 去的自然变换.有穷陨空间 X 的重心派生为一有穷陨空间,并且从它到 X 的自然变换是连续的.

从任意一个有穷陨空间 X_0 出发,我们可以按照下述的方式来构成一个谱.首先关于自然数 n 归纳地定义出一系列有穷陨空间 $X_n(n=0,1,\cdots)$,令 X_{n+1}为 X_n 的重心派生.又令

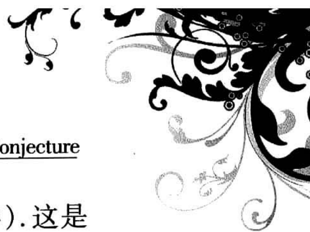
$$\pi_m^n=\pi_{m+1}\pi_{m+2}\cdots\pi_n$$

其中 π_k 为自 X_k 到 X_{k-1}的自然变换;π_m^n 则是自 X_n 到 $X_m(0\leqslant m\leqslant n)$的连续变换.容易看出,诸空间 X_n 与诸变换 π_m^n 合并而考虑,组成一个豪斯道夫谱.这样所构成的谱称为重心谱.特别,取 X_0 为 n – 维间架所有不定下集的全体在自然次序之下所成的陨空间,我们得到 n – 维元谱.П.С. Александров[3,19]证明了下面的定理:

n – 维元谱的低限为 n – 元体.

① 这个条件之并非必需,可以由下面的简例看出,设 Ξ 为按照自然次序的全体自然数所成的集.X_α 对于任意$\alpha\in\Xi$均为有穷陨空间$\{a,b,c\}$,其开集为:Λ,$\{a\}$,$\{a,b\}$,$\{a,c\}$,$\{a,b,c\}$;π_β^α为自 X_β 到 X_α 上的全同变换($\alpha<\beta(\Xi)$).不难看出,在这里 П.С. Александров 的条件并未满足,而谱的低限是一个由两个点组成豪斯道夫空间.





42. n - 元谱的特殊情形之一是所谓 n - 维单纯形(简称 n - 单纯形). 这是指包含在 r - 维欧几里得空间($r \geqslant n$),而不包含在它的任何维数小于 n 的下空间之内的 n - 维间架的凸闭包. 若间架 G 的凸闭包是单纯形 S, 则 G 唯一地决定了单纯形 S. G 的元称为 S 的顶点. G 的不空下集也定义单纯形,称为 S 的面. S 的面之中有非真面(即单纯形 S 自己). S 其他的面称为真面. S 所有的真面之和称为 S 的边缘.

设 C 为 r - 维欧几里得空间 E^r 中(不同维数的)单纯形集. 假如下面的条件适合,我们称 C 为 E^r 中的一个欧几里得单纯复合形(简称 E - 复合形):

(1) C 中每一个单纯形的面仍然属于 C;

(2)属于 C 的任意两个单纯形,其交或为空,或为它们的公共面.

这时,若 C 中含有 n - 维单纯形,而不含有 m - 维单纯形当 $m > n$,则称 C 为 n - 维复合形.

E^r 中的 E - 复合形 C 的所有单纯形之和,成为 E^r 中的一个点集. 若 E^r 中的下集 A 是 E^r 中某一 E - 复合形中诸单纯形之和,则称 A 为 E^r 中的直多面体. 特别若 A 为一 n - 维 E - 复合形中诸单纯形之和,则称 A 为 n - 维直多面体. 容易证明,以这种方式对于所有的直多面体都定义的初等几何维数由直多面体一义地决定. 换言之,同一个直多面体不可能同时是 n - 维而又是 m - 维的, $m \neq n$.

所有的同胚于直多面体的空间称为多面体,由 П.С.Александров 的结果得知,多面体为同胚于重心谱低限的空间.

每一个多面体可以同胚于许多直多面体,它们都具有同一个初等几何维数却是一件很不显然的事,实际上,这个事实的建立是 П.С.Урысон 与 K.Menger 所起始的维数的拓扑理论中一个主要的结果. 现在我们转到对它的简短叙述.

43. 布劳威尔-Урысон-Menger 所给的维数的拓扑定义,其基础的几何思想是庞加莱在 1912 年说出而最早的数学化陈述是 L.E.J. 布劳威尔在 1913 年给出的. 从改良了的布劳威尔维数定义出发, П.С.Урысон[3,6] 与 K. Menger 互相平行而独立地创造了维数理论深刻与丰富的内容. 这个理论的应用范围最初是限于列紧体,后来为 Л.А.Тумаркин[1,2,4] 及别的人推广到任意可数重的计量空间(即可数重的常空间)①. 以下直到标号 49,我们提到"空间"时,都是指可数重的计量空间.

Урысон-Menger 的维数定义是归纳的定义:维数小于等于 n 的空间类 ϑ_n 是对 n 归纳地定义起来的,自 $n = -1$ 开始. 我们现在给出这个定义比较简单的 menger 形式.

① 后来, П.С.Александров 把维数论的许多定理推广及于紧体(例如,见 П.С.Александров[13,15]).

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

定义 ϑ_{-1}只包含有空空间,仅此一个而已.设空间类 ϑ_{n-1}已定义就绪.若空间 X 的点 x 的任意近傍包含有一近傍,其边缘①属于 ϑ_{n-1},则称 X 在 x 点具有维数小于等于 n.若 X 在它的每一点的维数都小于等于 n,则称 X 具有维数小于等于 n.维数小于等于的空间类记为 ϑ_n.

显然,$\vartheta_{n-1}\subset\vartheta_n(n=0,1,2,\cdots)$.若 $X\in\vartheta_n\setminus\vartheta_{n-1}$,则称 X 为 n-维或具有维数 n.若 X 不属于任何 ϑ_n,则称 X 为无穷维.称 X 在 x 点为 n 维,假若 X 在此点具有维数小于等于 n 而不具有维数小于等于 $n-1$.空间 X 的维数记为 $\dim X$;X 在 x 点的维数记作 $\dim_x X$.

容易看出,所有的类 ϑ_n 均具有遗传性.换言之,若 X 为 Y 的下空间,则 $\dim X\leqslant\dim Y$.

维数论中,一个比较深刻的问题是关于和集的维数与各项的维数之间的关系.П.С.Урысон、Л.А.Тумаркин 及 menger 建立了下面关于求和的定理:

若空间可以表示为可数个维数小于等于 n 的闭集之和,则空间的维数小于等于 n.

对于不闭集的求和,情况就大不同.两个下空间之和的维数可以大于它们任意一个的维数;例如在直线上,有理数所成的0-维集与无理数所成的0-维集,二者之和是一个1-维空间——整个直线.设 A 与 B 是空间任意的下集,则它们的维数,照 П.С.Урысон 证明,应适合

$$\dim(A\cup B)\leqslant\dim A+\dim B+1$$

并且这个估值是不可再改良了.

与此密切相关的有下面可注意的关于分解的 Урысон 定理:

任何一个 n-维空间可以表示为 $n+1$ 个0-维下集之和,但不能表为少于$n+1$个0-维下集之和.

研究过的还有二空间乘积的维数与二空间维数之间的关系.容易证明

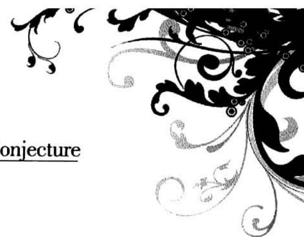
$$\dim(A\times B)\leqslant\dim A+\dim B \quad ②$$

此处"$A\times B$"表示空间 A 与 B 的乘积.曾经在某一段时期存在过这样的臆测,人们猜想式中的"$\leqslant$"号可以改为"$=$"号,但 П.С.Понтрягин[1]举出两个 E^4 中的列紧体 A 与 B,满足 $\dim A=2,\dim B=2,\dim(A\times B)=3$.Л.С.Понтрягин 的这个例子与后来 П.С.Александров 的"同调"维数论密切有关.

44.由②立刻知道 $\dim I^n\leqslant n$.证明此处等号成立是拓扑学的一大成就.这是由于好几个作者的工作而实现的,即由于布劳威尔、H·勒贝格、П.С.Урысон以及 K. Menger.这些工作中所证明的等式"$\dim I^n=n$",一方面说明 П.С.Урысон及 K.Menger 的维数观念是定义得很适当的,另一方面说明,在初

① 集 A 在空间 X 中的边线,定义为 A 在 X 中的闭包与 $X\setminus A$ 在 X 中闭包之交.





等几何意义之下的“维数”这个观念是一个拓扑不变量.

证明“dim $I^n = n$”这个等式的基础,第一是在于布劳威尔 - Lebesgue 关于 n - 维方体的遮的著名定理,第二是在于 П.С.Урысон[3,6]与 K.Menger 证明了上述归纳定义的维数与非归纳而与空间之遮有关的维数定义为同值.这个同值最初仅对于列紧体有了证明,在列紧体中,第二个维数定义是可以用计量的术语陈述的,后来,W.Hurewiez 把非归纳的定义推广到非列紧空间,我们此地介绍这个广义定义.

设 α 与 β 为空间 X 的遮,若 α 的每一元均包含于 β 的某一元之内,则称 α 为 β 的一个精炼.所谓遮 α 的阶是指满足下列条件的最大数 n:存在有互不相同的 α 中的 $n+1$ 个元,它们的交为不空.

维数的第二个定义是这样说的,非空空间的任意有穷遮若都有一个阶数小于等于 n 的有穷精炼,而存在有穷遮,不具阶数小于 n 的有穷精炼时,空间称为 n - 维.若对于任何自然数 n,某不空空间都不是 n - 维时,则称之为无穷维空间.空空间的维数定为 - 1.

П.С.Урысон[3,6]、Menger 及 Hurewicz 证明,第二个维数定义与归纳的定义为同值.

勒贝格 - 布劳威尔定理说,在第二个定义之下 I^n 为 n - 维.由同值定理立刻就得到 dim $I^n = n$.

由这个等式就知道,n - 维元体的维数是 n,并且所有的直多面体,在初等几何意义之下右是 n - 维,则其维数为 n,直多面体在初等几何意义之下的维数,因此是一个拓扑不变量,上面我们定义“多面体”为与直多面体同胚的空间,现在我们知道,与某一已知多面体 X 为同胚的诸直多面体具有相同的初等几何维数,等于 dim X.

45.上面引到的 Lebesgue - 布劳威尔定理与另一个著名的拓扑学定理密切有关,与布劳威尔不动点定理有关,这个定理说,自 n - 维元体到它自己的任意一个连续变换,必定至少把某一点变到自己.这个定理被 A.H.Тихонов 作了可观的推广,推广到 n - 维元体的无穷维类似.

A.H.Тихонов 考虑所谓线性拓扑空间,就是一个阿贝尔拓扑群,在其中定义了元与实数的乘法满足普通线性空间所满足的公理,并且 λx 连续地倚赖于元对 $<\lambda, x>$,此处 λ 为一实数,x 为群中一元.在这种空间中很自然地可以定义“线段”及“凸集”的观念.若线性拓扑空间每一个的近傍都包含有一个此点的凸近傍,则称之为局部凸线性拓扑群.

A.H.Тихонов 的定理说:

若 X 为一局部凸线性拓扑空间内的凸紧体,则自 X 到它自己的任意一个连续变换,必定至少把某一点变到自己.

这个定理有许多有趣的应用.A.A.Тихонов 自己利用它来证明无穷系常微分方程式解的存在.A.A.Марков[1]利用它来证明,关于"阿贝尔集"为不变的"均值"存在定理,后者又被用来证明"不变测度"与"积分不变量"的存在[5,7].A.Weil①也用了 A.H.Тихонов 的定理.这样,近代拓扑学里的一个中心思想就在其他的数学分支当中引出了丰富的结论.

46.在维数论中更进一步的巨大成就为 П.С.Александров[67]所得到,他证明了著名的关于以同样维数的多面体去"逼近"列紧体的定理.

这类定理之一是关于包含在 r - 维欧几里得空间 E^r 中或希尔伯特空间 H 中的列紧体.设 f 为自 E^r(或 H)中的集 X 到 E^r(或 H)中的连续变换.若 $\rho(x, fx) < \varepsilon$ 对于所有的 $x \in X$ 成立,则称 f 为 ε - 位移.此处 ρ 为 E^r(或 H)中的欧几里得(或希尔伯特)计量,ε 为正数.

П.С.Александров 关于 ε - 位移的定理说。

设 X 为 E^r(或 H)中的 n - 维列紧体,则对于任意 $\varepsilon > 0$ 必存在 ε - 位移将 X 变到一个 n - 维直多面体上去,并且存在 $\varepsilon > 0$ 使得 X 不可能经过 ε - 位移变换到一个维数较低的直多面体上去.

在另一个定理中牵涉列紧体的"ε - 变换".设 f 为自计量空间 X 到空间 Y 的连续变换,$\varepsilon > 0$.若 $f(x)$ 的反象所成的集,对所有的 x,均为直径小于 ε 的集,则 f 称为一个 ε - 变换②.

П.С.Александров 关于 ε - 变换的定理说:

若 X 为 n - 维列紧体,则对任何 $\varepsilon > 0$ 必存在 ε - 变换将 X 变到 n - 维多面体上去,并且存在 $\varepsilon > 0$ 使得 X 不可能经过一个 ε - 变换变到较低维的多面体上去.

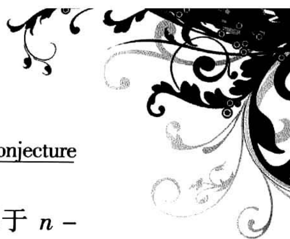
显然,这两个定理都可以作为 n - 维列紧体的特征性质.后来,关于 ε - 变换的定理被推广到具有可数重的常空间,在它的陈述中所用到的计量性质完全被消除③.ε - 变换的地位为"α - 变换"所代替,此处 α 是空间的一个有穷遮.设 f 是自空间 X 到空间 Y 的连续变换,α 为 X 的有穷遮,若 Y 中任意一点都有一近傍,其 f - 反象包含于 α 的某元之内,则称 f 为 α - 变换.则有下面对于 П.С.Александров定理的推广:

设 $\dim X = n$,α 为 X 的任意一个有穷遮,则必存在 α - 变换将 X 变到一个 n - 维多面体上去;并且存在 X 的有穷遮 α 使得 X 不可能经过一个 α - 变换变到较低维数的多面体上去.

47.维数还有一个重要的特征也是 П.С.Александров 所建立的,这个特征

① A. Weil. L'intégration dans les groupes topologigues.
② 计量空间 X 中的下集 A 的直径是定义为 A 中两点间距离的上界.
③ 见 W.Hurewicz and H. Wallman, Dimension Theory, 1941.





关联到自列紧体到 n - 维元体的连续变换.任何一个 n - 维元体都同胚于 n - 维单纯形,由布劳威尔关于区域及边缘的拓扑不变性的著名定理立刻知道,自单纯形 X 到元体 Y 的任何一个拓扑变换必将单纯形的边缘变到 Y 的同一个下集.这个下集称为元体 Y 的边缘. $n+1$ 维元体的边缘显然同胚于 n - 维球,就是 E^{n+1}中由下列关系所定的下集

$$\sum_{i=1}^{n+1} x_i^2 = 1$$

其中,$x_i(i=1,\cdots,n+1)$是坐标.

设 f 为自空间 X 到 n - 维元体 Y 上的连续变换,Y 的边缘为 S.假若不存在自 X 到 S 的连续变换 g,满足 $gx=fx$ 对于所有 $fx\in S$ 的点 x 成立,则称变换 f 是本质的.有下面的 П.С.Александров 定理:

自空间 X 到 n - 维元体有本质变换存在的充要条件是 $\dim X\geqslant n$.

这个定理首先为 П.С.Александров[7]对于列紧体证明;后来被成功地推广到任意具可数重的常空间上去①.事实上,最近 П.С.Александров 成功地把这个定理推广及于更广义的空间类,关于空间之重的限制完全取消了.

П.С.Александров 的这个定理在维数论中担任一个重要角色,它是连接 Урысон-Menger 维数论与 П.С.Александров 的同调维数论的中心环节②.

48.Л.С.Понтрягин 与 Л.Г.Шнирелвман[1]成功地把列紧体的维数与空间的计量性质联系起来.

每一个可计量空间允许有许多不同的计量,就是说,照前面的定义,对应了许多不同的计量空间,设 ρ 为列紧体 X 的某一计量,ε 为正数,则 X 可以用不同的方式表示为有穷多个在计量 ρ 之下直径小于 ε 的闭集之和.所必需的这种闭集的最少数目是与列紧体 X,与它的度量 ρ 以及正数 ε 有关.这个数目记为 $N(X,\rho,\varepsilon)$.

Л.С.Понтрягин 与 Л.Г.Шнирелвман 定理可表示为等式

$$\inf_{\rho} \lim \inf_{\varepsilon\to 0}\left(-\frac{\log N(X,\rho,\varepsilon)}{\log \varepsilon}\right) = \dim X$$

此处 inf lim 是关于列紧体 X 所有的计量 ρ 而取;lim inf 是对于一个固定的 ρ 关于 $\varepsilon\to 0$ 而取.

49.某些主要结果与“嘉当流形”的观念有关.若 X 为一 n - 维通列紧体,满足下面的条件则称为 n - 维嘉当流形:X 不可能表示为两个有别于 X 的闭集 A 与 B 的和,$\dim(A\cap B)<n-1$. Л.А.Тумаркин[3]及 Hurewicz 同时独立地证明了下面重要而深刻的定理:

---

① 见前面所举的 Hurewicz 与 Wallman 的群,其中 П.С.Александров 定理以略有改变的形式出现,未援引作者.

② 下面我们将讲到同调的维数论.

任何一个 n - 维列紧体必包含有 n - 维嘉当流形.

П.С.Александров[4]证明了下面关于 n - 维欧几里得空间 E^n 之下集的结果:

若包含于 E^n 中的列紧体 X 是二不相交通开集的境界,则 X 是 $n-1$ 维嘉当流形.

当 $n \leqslant 3$ 时,这个定理早已为 Урысон[14]证明.

50.庞加莱所首倡的代数方法在拓扑学中担任一个非常重要的角色.苏维埃拓扑学派的巨大成就是和这种在总名称"同调与交截的理论"之下的方法的发展与应用有关联的.特别是在这上面建立了 П.С.Александров 的"同调维数论",前面已经顺便提到过.Л.С.Понтрягин 的"对偶定律"与 А.Н.Колмогоров-Alexander 的"∇ - 同调"论也是与同调及交截的理论有关的,这些都是研究的有力武器,在拓扑学本身领域之内以及在其他交错的区域之内得出许多有价值的结果.现在我们转到对"同调论"基本情况的叙述.

这个理论的特征就是吸收阿贝尔群进来以构成空间的拓扑不变量,由此而产生的对象——"同调群"不仅与所考虑的空间有关,并且与引入的阿贝尔群,所谓系数域有关①.系数域可以由任何阿贝尔群充任,或为"陨"群,即未经拓扑化的,或为拓扑群.引入拓扑群作为系数域是 Л.С.Понтрягин[2,4]的成果,这使得有可能将"对偶定律"经过适当的途径建立在 Л.С.Понтрягин[3]"指标"论的基础之上.

以下所碰到的阿贝尔群都以"加法来书写",就是说,群运算以"+"号来表示,元 x 的逆元以"$-x$"来表示,群的单位元以"0"表示.所谓"群",了解为或是陨阿贝尔群或是紧阿贝尔拓扑群,ϑ 表示所有整数的全体所成的加法群,ϑ_m 表示整数 mod $m(m=2,3,\cdots)$的同余类所成的群,R 表示实数 mod 1 所成的群.所有这些群除了最后一个都看做陨群;R 看做自然拓扑化了的拓扑群.R 是一个紧群,它的空间同胚于圆周.群 ϑ_m 有时也看做紧拓扑群.

所谓陨群的指标,是指由它到 R 的一个准同构;所谓紧群的指标,依照 Л.С.Понтрягин,是指由它到 R 的一个连续准同构②.群 Γ 的指标可以按照下面的公式来相加

$$(\varphi+\psi)x=\varphi x+\psi x, \quad x\in\Gamma$$

此时 φ,ψ 是待相加的指标,右方表示在 R 中的相加.Γ 任意两个指标相加,所得的仍然是 Γ 的指标,容易看出,Γ 的全体指标在加法之下形成了一个阿贝尔群.这个群称为群 Γ 的指标群,记作 $\chi\Gamma$.这时,陨群 Γ 的指标群可以按照下述

① 不仅如此,还依赖于其他的因素,我们将来要提到.

② Л.С.Понтрягин 的指标论仅仅是关于可数陨群与具可数重列紧拓扑群间的对偶.不难将这理论推广到关于任意陨群与紧拓扑群间的对偶,如 van Kampen 所为.





的方式予以拓扑化.群 $\chi\Gamma$ 的下集 A 看做 $\chi\Gamma$ 中的开集,假如它满足下列的条件:对于 A 中每一元 φ 可以找到有穷个 Γ 中的元 $x_1,\cdots,x_k$ 及正数 δ 使得所有满足

$$|\varphi x_i - \psi x_i| < \delta(\bmod 1), \quad i=1,\cdots,k$$

的 Γ 的指标 ψ 都属于 A.陨群的指标群这样拓扑化之后是一个紧拓扑群.

Л.С.Понтрягин 证明,群和它的指标群之间的关系在某种意义下为互相可逆的;他证明,紧拓扑群 G 为某一陨群的拓扑化指标群的充要条件是 Γ 与 χG 同构.换言之,指标群的建造,定义了同构陨群所成的类与同构紧拓扑群所成的类之间一个 1-1 对应.这个陨群与紧群之间的"对偶"对应还可再进一步发展.例如,对于每一个自陨群 Γ 到陨群 ϑ 的准同构 h 对应了一个共轭的连续准同构 g,把紧群 $\chi\vartheta$ 变到紧群 $x\Gamma$,由下式来定义,即

$$(g\varphi)x=\varphi(hx), \quad x\in\Gamma,\varphi\in\chi\vartheta$$

这个在准同构之间建立起来的对应也是互相可逆的,逆对应的定义与前面类似.

我们说,群 G 对偶于群 Γ,假如 G(拓扑)同构于 Γ 的指标群.

Л.С.Понтрягин 对于群建立了"谱论"①,考虑了下面两种互为对偶的结构:

(1)设对于有向集 Ξ 的每一元 α 对应了一个紧拓扑群 Γ_α;对于每一对元 $<\alpha,\beta>$ 适合 $\alpha<\beta(\Xi)$ 的,对应了一个连续准同构 ω_α^β 自群 Γ_β 到群 Γ_α,而且

$$\omega_\alpha^\beta\omega_\beta^\gamma=\omega_\alpha^\gamma$$

只要 $\alpha<\beta<\gamma(\Xi)$.群系 Γ_α 及准同构系 ω_α^β 合并考虑称为一个倒谱.

同前面所考虑的空间谱备限观念相似,可以介绍倒谱的极限的观念.由群 $\Gamma_\alpha(\alpha\in\Xi)$ 及准同构 $\omega_\alpha^\beta(\alpha<\beta(\Xi))$ 所形成的倒谱,其极限是定义为由条所形成的拓扑群;所谓条就是元系 $<x_\alpha>_{\alpha\in\Xi}$ 适合 $x_\alpha\in\Gamma_\alpha(\alpha\in\Xi)$, $x_\alpha=\omega_\alpha^\beta x_\beta(\alpha<\beta(\Xi))$.两个条的加法由下列公式定义

$$<x_\alpha>+<y_\alpha>=<x_\alpha+y_\alpha>$$

而由条所成的空间,其拓扑结构以下述方式定义.条集 A 称为开集,假如它满足下列条件:若 $<x_\alpha>\in A$,则存在 $\beta\in\Xi$ 及拓扑群 Γ_β 中的开集 U_β,使得所有的条 $<y_\alpha>$,其中 $y_\beta\in U_\beta$ 也 $\in A$.不难看出,条的全体关于加法形成一个阿贝尔群,条的相加及其逆算——相减,关于所定的拓扑结构为连续的.换言之,倒谱的极限是定义为一阿贝尔拓扑群.不难看出,倒谱的极限是紧拓扑群.

(2)第二种结构是关于陨群的.设对于有向集 Ξ 的任一元对应了一个陨群

---

① 这个理论最初仅仅关于可数陨群与具可数重列紧拓扑群(见前一个脚注).N.E.斯廷罗德把这个理论推广到任意陨群与任意紧拓扑群.

Γ_α,对于任意一对元 $<\alpha,\beta>$ 适合 $\alpha<\beta(\Xi)$,对应了一个自群 Γ_α 到群 Γ_β 中的准同构 π_β^α.设这些群 Γ_α 都是互不相交的,并且准同构 π_β^α 满足

$$\pi_\gamma^\beta\pi_\beta^\alpha=\pi_\gamma^\alpha(\alpha<\beta<\gamma(\Xi))$$

群系 Γ_α 以及准同构系 π_β^α 合并起来考虑称为一个顺谱.

现在来看由群 $\Gamma_\alpha(\alpha\in(\Xi))$ 及准同构 $\pi_\beta^\alpha(\alpha<\beta(\Xi))$ 所形成的顺谱.诸群 Γ_α 的集合论的和集 S 之中可定义半序;令 $x<y(S)$,假如存在 $\alpha,\beta\in\Xi$,使得 $\pi_\beta^\alpha x=y$. S 中的元 x,y 称为同值,假如有一元 z 存在使得 $x<z(S),y<z(S)$.不难看出,这样定义的同值关系在 S 中是反射的,对称的,可递的.因此就把 S 分成了许多同值类.每一个这样的类称为所考虑的顺谱的一个束. S 中的每一元属于且仅属于一个束.

不难看出,对于顺谱当中任意两个束 u 与 v,必定可以找到群 Γ_α 使得 $u\cap\Gamma_\alpha\neq\Lambda$ 及 $v\cap\Gamma_\alpha\neq\Lambda$.这样就使得有可能按照下述的方式来定义束的加法.已给两个束 u 与 v,选取 α 使得 $u\cap\Gamma_\alpha\neq\Lambda$ 及 $v\cap\Gamma_\alpha\neq\Lambda$.设 $u\cap\Gamma_\alpha=x,v\cap\Gamma_\alpha=y$.将 x 与 y 在 Γ_α 中相加,所得到的群 Γ_α 中的元,属于且只属于一个束 w,易知这个束与 α,x 与 y 的选择无关.置 $u+v=w$.在这个定义之下,束的全体形成一个阿贝尔群.这个阿贝尔群称为所考虑的顺谱的极限.

顺谱和倒谱在下述的意义之下是互为对偶的.设已给一个倒谱,由紧拓扑群 $\Gamma_\alpha(\alpha\in\Xi)$ 及连续准同构 $\omega_\alpha^\beta(\alpha<\beta(\Xi))$ 所形成.看这些群的指标群 $\chi\Gamma_\alpha$.设 π_β^α 表示自群 $\chi\Gamma_\alpha$ 到群 $\chi\Gamma_\beta$,对偶于 ω_α^β 的准同构(见前面).则群 $\chi\Gamma_\alpha$ 与准同构 π_β^α 合并起来考虑,形成了一个顺谱,这个顺谱的极限同构于原来倒谱极限的指标群.类似的做法可以从一个已给的顺谱达到与它对偶的倒谱.

51.同调论最简单的对象是 A - 复合形.我们将考虑一个固定的 A - 复合形 C.设 Γ 为任意一个群,r 为一非负整数,考虑函数,其变数是复合形 C 的 $r+1$ 个顶点,其值是 Γ 中元.这样的函数 f 如果满足下列条件,则称为 C 中的一个 $<r,\Gamma>$ - 链:

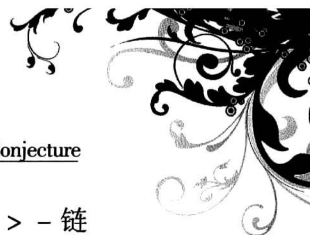
(1)$f(x_0,\cdots,x_r)\neq0$ 仅当 $x_0,\cdots,x_r$ 是复合形 C 某一间架的不同顶点时成立;

(2)$f(x_0,\cdots,x_{i-1},x_i,\cdots,x_r)=-f(x_0,\cdots,x_i,x_{i-1},\cdots,x_r),0<i\leqslant r$.当 $r=-1$,我们也定义 C 的 $<r,\Gamma>$ 链.群 Γ 中的元称为 C 中的 $<-1,\Gamma>$ - 链(这就相当于把具有零个变数的函数看做常数,后来将较为方便).条件(1)当 $r=0$ 时当然满足,这时 f 为只含一个变数的函数. C 中的 $<r,\Gamma>$ - 链自然地可以相加

$$(f+g)(x_0,\cdots,x_r)=f(x_0,\cdots,x_r)+g(x_0,\cdots,x_r) \quad ④$$

显然,C 中二 $<r,\Gamma>$ - 链之和仍然是一个 C 中的 $<r,\Gamma>$ - 链.很清楚,C 中





的 $<r,\Gamma>$ - 链关于加法形成一阿贝尔群. 这个群——C 中的 $<r,\Gamma>$ - 链群——将记为 $L^r(C,\Gamma)$.

若此时 Γ 为一拓扑群,则在 $L^r(C,\Gamma)$中可以按照下述方式定义拓扑结构. 这个群的下集 A 定义为开集,假如对于 A 中的每一元f 可以找到群 Γ 的零元的一个近傍 U,使得凡是使关系

$$f(x_0,\cdots,x_r)-g(x_0,\cdots,x_r)\in U$$

对于所有的 $x_1,\cdots,x_r$ 都成立的 $<r,\Gamma>$ - 链 g,必定属于 A. 若 Γ 为紧群,则 $L^r(C,\Gamma)$为紧群.

在同调论中,两个"边缘"算子占据着基本的地位,一个是"下","古典"的边缘算子,通常记作"$\triangle$",另一个是比较后期才介绍进来的"上边缘算子",记作"∇". 这两个算子都是作用于 C 中的 $<r,\Gamma>$ - 链的,下边缘仅限于 $r\geqslant 0$,上边缘无任何限制. 下边缘算子作用于 $<r,\Gamma>$ - 链 $f(r\geqslant 0)$得出一个 $<r-1,\Gamma>$ - 链$\triangle f$,按照下列公式定义

$$\triangle f(x_0,\cdots,x_{r-1})=\sum_x f(x,x_0,\cdots,x_{r-1}) \qquad ⑤$$

这里的和是展开于复合形 C 所有的顶点 x 之上的. 上边缘算子作用于 $<r,\Gamma>$ - 链 f 得出一个 $<r+1,\Gamma>$ - 链∇f,由下式定义

$$\nabla f(x_0,\cdots,x_{r+1})=\begin{cases}\sum_{k=0}^{r+1}(-1)^k f(x_0,\cdots,x_{k-1},x_{k+1},\cdots,x_r) & ,\text{对于}\{x_0,\cdots,x_{r+1}\}\in C\\ 0 & ,\text{其他情形}\end{cases}$$

两个算子对于加法都适合分配律,而且

$$\triangle\triangle f=0,\nabla\nabla f=0 \qquad ⑥$$

由于可分配性,算子$\triangle$定义了一个自群 $L^r(C,\Gamma)$到群 $L^{r-1}(C,\Gamma)$中的准同构,而算子∇定义了一个 $L^r(C,\Gamma)$到 $L^{r+1}(C,\Gamma)$的准同构. 若 Γ 为拓扑群,则这些准同构是连续的. 第一个准同构的核是 $L^r(C,\Gamma)$的一个下群,由 C 中所有适合 $\triangle f=0$ 的 $<r,\Gamma>$ - 链 f 所形成. 这些链称为 C 中关于 Γ 的 r - 维$\triangle$ - 闭链. 第二个准同构的核为 $L^r(C,\Gamma)$的一个下群由 C 中所有适合 $\nabla f=0$ 的 $<r,\Gamma>$ - 链 f 所形成. 这些链称为 C 中关于 Γ 的 r - 维 ∇ - 闭链. 第一个下群记作 $Z^r_\triangle(C,\Gamma)$,第二个记作 $Z^r_\nabla(C,\Gamma)$. 若 Γ 为拓扑群,则二者在 $L^r(C,\Gamma)$中都是闭集.

由于⑥,在$\triangle$所定义的准同构之下,$L^{r+1}(C,\Gamma)$的象为 $Z^r_\triangle(C,\Gamma)$的一个下群;在∇所定义的准同构之下,$L^{r-1}(C,\Gamma)$的象为 $Z^r_\triangle(C,\Gamma)$的一个下群. 若 Γ 为紧拓扑群,这些下群在 $L^r(C,\Gamma)$中都是闭的. 第一个记作 $H^r_\triangle(C,\Gamma)$,第二个记作 $H^r_\nabla(C,\Gamma)$. C 中属于$H^r_\triangle(C,\Gamma)$的 r - 维$\triangle$ - 闭链,谓之在 C 中$\triangle$ - 同调于零;C 中属于$H^r_\nabla(C,\Gamma)$的 r - 维 ∇ - 闭链,谓之在 C 中 ∇ - 同调于零. 按

照这样的定义，r－维$\triangle$－闭链 f 关于 Γ 在 C 中$\triangle$－同调于零的充要条件是在 C 中存在 $<r+1,\Gamma>$－链 g 满足 $f=\triangle g$；r－维∇－闭链 f 关于 Γ 在 C 中∇－同调于零的充要条件是在 C 中存在 $<r-1,\Gamma>$－链 g 满足 $f=\nabla g$.

因为 $H^r_\triangle(C,\Gamma)\subset Z^r_\triangle(C,\Gamma)$ 以及 $H^r_\nabla(C,\Gamma)\subset Z^r_\nabla(C,\Gamma)$，所以可以作商群 $Z^r_\triangle(C,\Gamma)/H^r_\triangle(C,\Gamma)$ 以及 $Z^r_\nabla(C,\Gamma)/H^r_\nabla(C,\Gamma)$. 若 Γ 为紧拓扑群，则这些群都可看做拓扑群，并且所有这些都是紧拓扑群. 群 $Z^r_\triangle(C,\Gamma)/H^r_\triangle(C,\Gamma)$ 及 $Z^r_\nabla(C,\Gamma)/H^r_\nabla(C,\Gamma)$ 分别记为 $B^r_\triangle(C,\Gamma)$ 及 $B^r_\nabla(C,\Gamma)$，称为复合形 C 关于 Γ 的 r－维贝蒂群，第一个称为贝蒂$\triangle$－群，第二个称为贝蒂∇－群. $B^r_\triangle(C,\Gamma)$ 的元是 C 中互为$\triangle$－同调的 C 中关于 Γ 的 r－维$\triangle$－闭链所形成的类，$B^r_\nabla(C,\Gamma)$ 的元是 C 中互为∇－同调的 C 中关于 Γ 的 r－维∇－闭链所形成的类. 这样，我们说 C 中两个 $<r,\Gamma>$－链在 C 中为$\triangle$－同调，假如它们的差在 C 中$\triangle$－同调于零，类似地定义二 $<r,\Gamma>$－链在 C 中∇－同调.

现在设 G 为 Γ 的指标群，则容易看出群 $L^r(C,\Gamma)$ 对偶于群 $L^r(C,\Gamma)$，并且算子$\triangle$所定义的自 $L^{r+1}(C,G)$ 到 $L^r(C,G)$ 的准同构，共轭于算子∇所定义的自 $L^r(C,\Gamma)$ 到 $L^{r+1}(C,\Gamma)$ 的准同构. 由此推出，群 $B^r_\triangle(C,\Gamma)$ 与群 $B^r_\nabla(C,\Gamma)$ 是对偶的.

52. 设 D 为复合形 C 的下复合形，即这个复合形的下集面本身也成一个复合形. 集 $C\setminus D$ 通常并不是复合形. 然而可以合理地介绍“$C\setminus D$ 中 $<r,\Gamma>$－链”的观念. 属于 $C\setminus D$ 的间架，其所含的元称为 $C\setminus D$ 的顶点. $C\setminus D$ 中 $<r,\Gamma>$－链的观念可以逐字逐句依照 C 中 $<r,\Gamma>$－链的观念来定义，只消到处把“C”易为“$C\setminus D$”$C\setminus D$ 的 $<r,\Gamma>$－链在加法之下形成一个群 $L^r(C\setminus D,\Gamma)$，加法是按照公式④来定义的. 对于 $\{x_0,\cdots,x_{r-1}\}\in C\setminus D$，下边缘算子由公式⑤定义，对于 $\{x_0,\cdots,x_{r-1}\}\in D$ 则含 $\triangle f\{x_0,\cdots,x_{r-1}\}=0$. 上边缘算子的定义如前，只需作相应的修改. $\triangle$和∇的可分配性仍然保持，等式⑥仍然有效. 但这已是贝蒂群的定义中所需要的全部. 重复前一节的做法，加以相应的修改，我们就得到贝蒂群 $B^r_\triangle(C\setminus D,\Gamma)$ 与 $B^r_\nabla(C\setminus D,\Gamma)$ 的定义. 若 Γ 对偶于 G，则如同对于群 $B^r_\triangle(C,\Gamma)$ 与 $B^r_\nabla(C,\Gamma)$ 一样可以建立群 $B^r_\triangle(C\setminus D,\Gamma)$ 与 $B^r_\nabla(C\setminus D,G)$ 之间的对偶关系.

53. 贝蒂群在拓扑学中的意义，首先是，作为拓扑不变量. 它可以对于很广义的空间定义. 对于多面体，贝蒂群很早就定义了，在庞加莱的工作中就已出现. 已给的多面体 X 必然是某直多面体 Y 在某拓扑变换之下的象；Y 是某一个 E－复合形 K 的全体单纯形之和；最后，K 又定义了一个 A－复合形 C，由属于 K 的单纯形的顶点间架所组成. 多面体 X 的贝蒂群就定义为 A－复合形 C 的贝蒂群. 这个定义的合理性是很不显然的，因为其中包含着几个任意的选择：直多面体 Y 及 E－复合形 K 的选择可以是多种多样的. 但是，美国数学家亚历山





大证明,这样所得到的贝蒂群仅仅与多面体 X 有关,这同时证明了贝蒂群的拓扑不变性.虽然亚历山大所涉及的仅仅是关于某些特殊系数域的贝蒂群,并且只涉及△-群,但这种限制并无原则上的意义.

其后,无论是在苏联或是在外国,有许多拓扑学者致力于以一种适当的方式推广贝蒂群的定义到尽可能广义的空间类上去.所谓“Alexander-Понтрягин 对偶定律”是对各种尝试性定义的试金石.这个定理建立 n-维欧几里得空间 E^n 中下集的贝蒂群与其余集的贝蒂群之间的关系.亚历山大最初证明这个定律是关于 E^n 中的多面体及其余集,正当他定义了 E^n 中开集的贝蒂群之后.第一个尝试对于列紧体定义贝蒂群的是 L. Vietoris,在 1927 年.但是,这个定义不是十分成功的,因为在这个定义之下得不出“对偶定律”,Л. С. Понтрягин[2,4]得到了实质上巨大的成就,他引入了拓扑群作为系数域并且应用了他关于指标的理论,这使得他有可能建立亚历山大型的“对偶定律”,E^n 中列紧体关于系数域的贝蒂群与其余集关于对偶的可数陨系数域的贝蒂群之间的关联得以建立.

后来 А. Н. Колмогоров[1,2,3,4,5]作了重大推进,他第一个介绍了上边缘算子,并以任意局部紧空间给了△-和▽-贝蒂群的定义①.这些群的非常广义的定义,斯廷罗德也会给过.但这个定义并不是很完满的:对于非常简单的局部紧空间如像直线,这个定义给出了极为复杂的贝蒂群.进一步作了本质上改良的是 П. С. Александров[14,16,17,18],对于法空间的开下集定义了“内贝蒂群”,并对于这些群证明了“Колмогоров 对偶定律”.这个利用顺谱及倒谱来定义的“内贝蒂群”免去了 Betti-Колмогоров 斯廷罗德群的缺点,如同 Г. С. Чогошвили[3]所证明,内贝蒂群与 Betti-Колмогоров 群相同.我们现在转到对这个群的理论的基本事项作详细的叙述.

54.在“内 Betti-Александров 群”的定义中,除了已经讲到过的观念以外,有两个辅助的观念占据非常重要的地位:就是“单纯变换”及“神经”(Nerve)的观念.

设 φ 为一变换,把 A-复合形 D 的顶点的全体变换到 A-复合形 C 的顶点的全体.若 D 中每一间架的 φ-象为 C 中一间架时,则称 φ 为一单纯变换.

每一个自 A-复合形 D 到 A-复合形 C 的单纯变换 φ 诱导了两个算子:算子 ρ_φ 作用于 D 中的链,算子 σ_φ 作用于 C 中的链.第一个算子将 D 中的 $<r,\Gamma>$-链 f 变为 C 中的 $<r,\Gamma>$-链 $\rho_\varphi f$,由下面的式子定义

$$\rho_\varphi f(y_0,\cdots,y_r) = \sum_{\varphi x_i = y_i} f(x_0,\cdots,x_r)$$

① 上边缘算子同时独立地为亚历山大所引入.

这里的和是展开于所有适合 $\varphi x_i = y_i(i=0,\cdots,r)$ 的诸元 $x_0,\cdots,x_r$. 第二个算子将 C 中的 $<r,\Gamma>$ – 链 f 变为 D 中的 $<r,\Gamma>$ – 链 $\sigma_\varphi f$, 由下面的式子定义

$$\sigma_\varphi f(x_0,\cdots,x_r)=\begin{cases}f(\varphi x_0,\cdots,\varphi x_r) & ,对于\{x_0,\cdots,x_r\}\in D\\ 0 & ,其他情形\end{cases}$$

为了使这些算子对于 $r=-1$ 也有意义, 置 $\rho_\varphi f=\sigma_\varphi f=f, f\in\Gamma$. 算子 ρ_φ 显然定义了一个自 $L^r(D,\Gamma)$ 到 $L^r(C,\Gamma)$ 的准同构, 算子 σ_φ 定义了一个自 $L^r(C,\Gamma)$ 到 $L^r(D,\Gamma)$ 的准同构. 若 Γ 为拓扑群, 则这些准同构是连续的. 若 G 为 Γ 的特征群, 则 σ_φ 所定义的自 $L^r(C,G)$ 到 $L^r(D,G)$ 的准同构, 共轭于 ρ_φ 所定义的自 $L^r(D,\Gamma)$ 到 $L^r(C,\Gamma)$ 的准同构. 算子 ρ_φ 与算子△可交换, 算子 σ_φ 与算子▽可交换. 因此, ρ_φ 诱导了一个自群 $B^r_\triangle(D,\Gamma)$ 到群 $B^r_\triangle(C,\Gamma)$ 的准同构, σ_φ 诱导了一个自群 $B^r_\nabla(C,\Gamma)$ 到群 $B^r_\nabla(D,\Gamma)$ 的准同构. 若 Γ 为拓扑群, 则这些准同构是连续的. 若 G 为 Γ 的指标群, 则 ρ_φ 所诱导的自 $B^r_\triangle(D,\Gamma)$ 到 $B^r_\triangle(C,\Gamma)$ 的准同构共轭于 σ 所诱导的自 $B^r_\nabla(C,G)$ 到 $B^r_\nabla(D,G)$ 的准同构.

占有特别地位的是那一种情形, 当 D 是 C 的下复合形, 而 φ 为 D 的全体顶点的全同变换. 这时, ρ_φ 与 σ_φ 所诱导的关于链群及贝蒂群的准同构, 称为自然准同构. 这时, 相应地也有由算子 E 与 J 所定义的自链群 $L^r(C\setminus D,\Gamma)$ 到链群 $L^r(C,\Gamma)$ 以及自链群 $L^r(C,\Gamma)$ 到链群 $L^r(C\setminus D,\Gamma)$ 的"自然准同构"其定义是: 对于 $f\in L^r(C\setminus D,\Gamma)$

$$Ef(x_0,\cdots,x_r)=\begin{cases}f(x_0,\cdots,x_r), 若 f(x_0,\cdots,x_r)有定义\\ 0, 若 x_i 是 C 中顶点, f(x_0,\cdots,x_r)无定义\end{cases}$$

对于 $f\in L^r(C,\Gamma)$

$$Jf(x_0,\cdots,x_r)=f(x_0,\cdots,x_r)$$

若 x_i 是 $C\setminus D$ 中顶点, E 与▽可交换, J 与△可交换, 由此可知, E 诱导了一个 $B^r_\nabla(C\setminus D,\Gamma)$ 到 $B^r_\nabla(C,\Gamma)$ 的准同构. J 诱导了一个自 $B^r_\triangle(C,\Gamma)$ 到 $B^r_\triangle(C\setminus D,\Gamma)$ 的准同构. 这些准同构也称为自然准同构.

55. 设 α 为一有限系不空集, 则 α 中的不空下集具有不空交的显然形成一个 A – 复合形. 这个 A – 复合形称为系 α 的神经. 按照这个定义 α 的神经中的顶点就是 α 的元, 间架 $\{A_0,\cdots,A_r\}$, $A_i\in\alpha(i=0,\cdots,r)$ 属于 α 的神经的充要条件是

$$\bigcap_{i=c}^{r}A_i\neq A$$

α 的神经记作 $N(\alpha)$.

56. 设 α 与 β 为空间 X 的两个有限遮. 若 α 为 β 的精炼, 则对于遮 α 的每一元∪取遮 β 中包含它的一个元与之对应, 我们就得到自 A – 复合形 $N(\alpha)$ 到





A - 复合形 $N(\beta)$的一个单纯变换.按照前面所说的,这个单纯变换诱导了自群 $B^r_\triangle(N(\alpha),\Gamma)$及群$B^r_\triangle(N(\beta),\Gamma)$分别到群 $B^r_\triangle(N(\beta),\Gamma)$及群 $B^r_\nabla(N(\alpha),\Gamma)$的准同构,并且当 Γ 为拓扑群时,这些准同构是连续的.在构成这些准同构的过程中涉及关于对应于 α 中的每一元选择一个包含它的β 中的元,这种选取的任意性仅仅是表面的:最后所得的准同构与这个任意的选取并无关系.这样,对应于空间 X 的每一对应 $<\alpha,\beta>$ 而 α 是β 的精炼的,就唯一地定义了自群 $B^r_\triangle(N(\alpha),\Gamma)$到群 $B^r_\triangle(N(\beta),\Gamma)$的一个准同构,以及自群$B^r_\nabla(N(\beta),\Gamma)$到群 $B^r_\nabla(N(\alpha),\Gamma)$的一个准同构.

现在,从所考虑的遮之中取出那样的元,它们的闭包是不紧的.设 α_0 与 β_0 分别表示遮 α 及β 中所有的元,其闭包在 X 中是不紧的.则 $N(\alpha_0)$与 $N(\beta_0)$将分别为 $N(\alpha)$及 $N(\beta)$的下复合形,而上面所定义的单纯变换把 $N(\alpha_0)$单纯地变换到 $N(\beta_0)$之中去.我们将考虑链群 $L^r(N(\alpha),\Gamma)$, $L^r(N(\beta),\Gamma)$, $L^r(N(\alpha)\setminus N(\alpha_0),\Gamma)$,$L^r(N(\beta)\setminus N(\beta_0),\Gamma)$.对于前两个群的元施行自然准同构 J,则把它们分别变换为后两个群的元;对于后两个群的元施行自然准同构 E,则把它们分别变换为前两个群的元.如同上面说的那样选择复合形$N(\alpha)$到 $N(\beta)$的单纯变换 φ,则有算子 ρ_φ 与σ_φ,分别定义了自$L^r(N(\alpha),\Gamma)$到 $L^r(N(\beta),\Gamma)$以及自 $L^r(N(\beta),\Gamma)$到 $L^r(N(\alpha),\Gamma)$的准同构.这就使得有可能定义算子 $\rho_{\varphi 1}$及 $\sigma_{\varphi 1}$,它们分别定义自 $L^r(N(\alpha)\setminus N(\alpha_0),\Gamma)$到$L^r(N(\beta)\setminus N(\beta_0),\Gamma)$的准同构,以及自 $L^r(N(\beta)\setminus N(\beta_0),\Gamma)$到$L^r(N(\alpha)\setminus N(\alpha_0),\Gamma)$的准同构,其定义是按照下面的式子

$$\rho_{\varphi 1}=J\rho_\varphi E,\sigma_{\varphi 1}=J\sigma_\varphi E$$

$\rho_{\varphi 1}$与$\triangle$可交换,$\sigma_{\varphi 1}$与∇可交换,由此可知,算子 $\rho_{\varphi 1}$与 $\sigma_{\varphi 1}$分别诱导了自群 $B^r_\triangle(N(\beta)\setminus N(\beta_0),\Gamma)$到 $B^r_\triangle(N(\beta)\setminus N(\beta_0),\Gamma)$以及自群 $B^r_\nabla(N(\beta)\setminus N(\beta_0),\Gamma)$到 $B^r_\nabla(N(\alpha)\setminus N(\alpha_0),\Gamma)$的准同构.这些准同构也和单纯变换 φ 的选取无关.因此,对于空间 X 的每一对遮$<\alpha,\beta>$,而 α 是β 的精炼的,就唯一地定义了自群 $B^r_\triangle(N(\alpha)\setminus N(\alpha_0),\Gamma)$到群 $B^r_\triangle(N(\beta)\setminus N(\beta_0),\Gamma)$的一个准同构,以及自群 $B^r_\nabla(N(\beta)\setminus N(\beta_0),\Gamma)$到群 $B^r_\nabla(N(\alpha)\setminus N(\alpha_0),\Gamma)$的一个准同构.第一个准同构记为 ω^α_β,第二个记为 π^β_α.

57.设 α 与β 为空间X 的有穷遮,若 α 为β 的精炼而β 不是α 的精炼,则称β 在α 之前.在这种意义之下,X 所有的有穷遮全体形成了一个半序集.若 X 为属于 T_1 类的无穷空间,则它所有的有穷遮全体所形成的半序集为有向集(П.С. Александров[14]),有鉴于将来我们所要考虑的只有这一种情形,空间 X 的有穷遮全体所形成的有向集记作Ξ.我们固定一个不负整数 r 及拓扑群Γ.集 Ξ 的每一元α 对应了群

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

$$B^r_{\triangle_\alpha} = B^r_\triangle(N(\alpha) \setminus N(\alpha_0), \Gamma)$$

则对于 $\beta < \alpha(\Xi)$就有上面所定义的连续准同构 ω_β^α 把群 $B^r_{\triangle_\alpha}$变换到 $B^r_{\triangle_\beta}$. 不难看出,ω_β^α 满足条件③. 这样,诸群 $B^r_{\triangle_\alpha}$与诸准同构 ω_β^α 合并在一起考虑,形成一个倒谱,由空间 X,整数 r 及群 Γ 唯一决定. П.С. Александров[14]称这个谱的极限为空间 X 关于系数域 Γ 的 r - 维内贝蒂△ - 群. 这是一个紧拓扑群,我们把它记作 $B^r_\triangle(X, \Gamma)$.

同理,令

$$B^r_{\nabla_\alpha} = B^r_\nabla(N(\alpha) \setminus N(\alpha_0), \Gamma)$$

其中,Γ 为陨群,可得到一个由诸群 $B^r_{\nabla_\alpha}$及诸准同构 π_α^β 所形成的顺谱. П.С. Александров[14]称这个谱的极限为空间 X 关于系数域 Γ 的 r 维内贝蒂▽ - 群. 这是一个陨群. 我们把它记作 $B^r_\triangle(X, \Gamma)$①.

容易证明,若 G 为紧拓扑群 Γ 的对偶,则群 $B^r_\nabla(X, G)$为 $B^r_\triangle(X, \Gamma)$的对偶.

58. 设 Y 为空间 X 的闭下集,$Z = X \setminus Y$. 把 Y, Z 也看做空间,则三个空间 X, Y 与 Z 各有其内贝蒂群. 情形颇与上面所说过的关于复合形 C 中的下复合形 D 的情形类似,那时我们有复合形 C, D 及 $C \setminus D$ 的贝蒂群. 那时定义了自群

$$B^r_\triangle(D, \Gamma), B^r_\nabla(C \setminus D, G), B^r_\nabla(C, G), B^r_\triangle(C, \Gamma)$$

分别得到群

$$B^r_\triangle(C, \Gamma), B^r_\nabla(C, G), B^r_\nabla(D, G), B^r_\triangle(C \setminus D, \Gamma)$$

内的"自然准同构",其中 Γ 与 G 可以是任意的群. 与此相仿, П.С. Александров[16]成功地定义了自群

$$B^r_\triangle(Y, \Gamma), B^r_\nabla(Z, G), B^r_\nabla(X, G), B^r_\triangle(X, \Gamma)$$

分别得到群

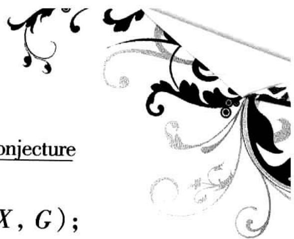
$$B^r_\triangle(X, \Gamma), B^r_\nabla(X, G), B^r_\nabla(Y, G), B^r_\triangle(Z, G)$$

内的"自然准同构",其中 Γ 为任意的紧群,G 为任意的陨群,X 为任意的局部紧法空间. "自然准同构"的构成是从复合形的情形利用谱来趋于极限而实现的. 这个构成步骤中的细节很繁复,我们不能在这样有限的篇幅内予以介绍.

由于"自然准同构"使 П. С. Александров 有可能定义下面诸群: $B^r_\triangle(XY, \Gamma)$——自然准同构之下 $B^r_\triangle(Y, \Gamma)$在 $B^r_\triangle(X, \Gamma)$中的象; $B^r_\triangle(Y{:}\,X, \Gamma)$——同上准同构的核; $B^r_\nabla(XZ, G)$——自然准同构之下 $B^r_\nabla(Z, G)$在 $B^r_\nabla(X, G)$中的象; $B^r_\nabla(Z{:}\,X, G)$——同上准同构的核; $B^r_\nabla(YX, G)$——自然准同构之下 $B^r_\nabla(X, G)$在 $B^r_\nabla(Y, G)$中的象; $B^r_\triangle(ZX, \Gamma)$——自炙准同构之下

① 在 П.С. Александров 的论文中,内贝蒂群仅对于包含在法空间内作为开集的空间定义.





$B^r_\triangle(X,\Gamma)$在 $B^r_\triangle(Z,\Gamma)$中的象；$B^r_\nabla(Y:X,G)=B^r_\triangle(Y,G)/B^r_\nabla(YX,G)$；$B^r_\triangle(Z:X,\Gamma)=B^r_\triangle(Z,\Gamma)/B^r(ZX,\Gamma)$. 自 $B^r_\nabla(X,G)$到 $B^r_\nabla(Y,G)$以及自 $B^r_\nabla(X,\Gamma)$到 $B^r_\triangle(Z,\Gamma)$的准同构的核并非是什么新的群：前者同于 $B^r_\nabla(XZ,G)$，后者同于 $B^r_\nabla(YX,G)$. П.С.Александров 称为 $B^r_\triangle(XY,\Gamma)$，$B^r_\nabla(XZ,G)$，$B^r_\triangle(Y:Y,\Gamma)$，$B^r_\nabla(Z:X,G)$，$B^r_\nabla(YX,G)$，$B^r_\nabla(ZX,\Gamma)$，$B^r_\nabla(Y:X,G)$，$B^r_\triangle(Z:X,\Gamma)$为空间组 X,Y,Z 的贝蒂群.

在群 G 是群 Γ 的对偶这一假设之下，П.С.Александров 对于空间组 X,Y,Z 的贝蒂群，证明了它们之间如下的关系

$$B^r_\triangle(XY,\Gamma)\mid B^r_\nabla(YX,G),B^r_\nabla(XZ,G)\mid B^r_\triangle(ZX,\Gamma),$$
$$B^r_\triangle(Y:X,\Gamma)\mid B^r_\nabla(Y:X,G),B^r_\nabla(Z:X,G)\mid B^r_\triangle(Z:X,\Gamma),$$
$$B^r_\triangle(Y:X,\Gamma)\mid B^{r+1}_\nabla(Z:X,G) \tag{7}$$
$$B^r_\triangle(Y:X,\Gamma)\approx B^{r+1}_\triangle(Z:X,\Gamma),B^r_\nabla(Y:X,G)\approx B^{r+1}_\nabla(Z:X,G)$$

其中“|”表示对偶，而符号“$\approx$”表示同构或拓扑同构. 在这里特别有趣与重要的是“主要对偶定理”⑦，这个定理联系相邻维数贝蒂群间的关系，其中之一是关于集 Y 的群，另一个关于 Y 在 X 中的余集(因此两个群都与其所对应的集在 X 中所处的位置有关).

由主要对偶定理可以得到对于“单通”空间情形的 Колмогоров-Alexander 对偶定律①. 我们说，空间 X 关于系数域 Γ 在维数 r 为单通的，假如

$$B^r_\triangle(X,\Gamma)=\{0\}$$

显然，在这个情形，$B^r_\triangle(Y:X,\Gamma)=B^r_\triangle(Y,\Gamma)$，$B^r_\nabla(Z:X,G)=B^r_\nabla(Z,G)$，其中 G 为 Γ 的指标群. 由此以及⑦就得到 Колмогоров-Alexander 对偶定律：若局部紧法空间 X 关于系数域 Γ 在维数 r 与 $r+1$ 是单通的，则对于它的任意闭下集 Y 有

$$B^r_\triangle(Y,\Gamma)\mid B^{r+1}_\nabla(X\setminus Y,G)$$

其中 G 为 Γ 的指标群(П.С.Александров[14,17]).

59. 利用空间组 X,Y,Z 的贝蒂群可以很简单地陈述 П.С.Александров“同调维数”观察的定义. 设 X 为局部紧法空间，Γ 为紧阿贝尔拓扑群. 若 n 为满足下列条件的正整数 r 中最大的一个，则称 X 关于 Γ 为 n - 维：X 包含有闭下集 Y，使得 $B^{r-1}_\triangle(Y:X,\Gamma)\neq\{0\}$；若这样的 r 不存在，则称 X 关于 Γ 为 0 - 维；若这样的 r 存在，但并无最大者，则称 X 关于 Γ 为无穷维. 这样，对每一个局部紧法空间 X 定义了关于 Γ 的维数. X 关于 Γ 的维数记作$\triangle(X,\Gamma)$.

在$\triangle(X,\Gamma)$的定义中不仅涉及空间 X，并且也涉及群

① 见 П.С.Александров[14,17].

Γ. Л.С.Понтрягин[1]造出 E^4 中列紧体 X 的实例，说明 $\triangle(X,\Gamma)$ 的确与 Γ 有关. 另一方面 Александров-Понтрягин-Франкль 有定理说：对于 E^4 中的列紧体 X，所有的数 $\triangle(X,\mathscr{G}_m)$，$m=2,3,4,\cdots$ 都与 $\dim X$ 相等(П.С.Александров[8]). 又证明，若 X 为多面体，则所有的数 $\triangle(X,\Gamma)$ 均与 $\dim X$ 相等(П.С.Александров[8]).

比较有原则上重要性的是 Александров-Понтрягин 定理，说：对于所有包含于有穷维欧几里得空间中的列紧体 X，有

$$\triangle(X,\mathscr{P})=\dim X$$

(П.С.Александров[8]，Л.С.Понтрягин[4]，П.С.Александров，Л.С.Понтрягин 与 H.Hopf[1]). 这样，如果限于 $E^n(n=1,2,\cdots)$ 中的列紧体，Урысон-Menger 维数就成为同调维数之一. 出了这个界限情况将是如何，直到现在还不知道. Александров-Понтрягин 定理的证明用到前面提过的 П.С.Александров 关于本质变换的定理，Hopf 关于自 n - 维多面体到 n - 维球的"本质变换"的同调特征化定理以及 Alexander-Понтрягин 对偶定理.

П.С.Александров[8]证明：对于所有 E^n 中的列紧体 X $\triangle(X,\Gamma)\leqslant\dim X$. 因此，如果限于这种列紧体，Урысон-Menger 维数就是同调维数中最大的一个.

维数的同调理论基本上是对于 E^n 中的列紧体以及对于系数域 $\mathscr{G}_m(m=2,3,\cdots)$，$\mathscr{P}$ 来建造的(对于有理数所成的加法群 $\mathscr{K}$，有理数 mod 1 的加法群 $\mathscr{R}_1$，及整数加法群 $\mathscr{G}$ 而言，$\triangle(X,\mathscr{K})$，$\triangle(X,\mathscr{K}_1)$，$\triangle(X,\mathscr{G})$ 也都有定义. 这些定义仍然在上面所述维数的同调理论的范围之内). 对于这些空间与系数域，许多类似于 Урысон-Menger 理论中定理的结果被证明. 特别是 П.С.Александров[8]对于同调维数论证明了求和的定理，证明了列紧体必包含有嘉当流形的定理. 它们的陈述都是与 Урысон-Menger 理论中相应的定理类似，只消把 $\dim X$ 换为 $\triangle(X,\Gamma)$，并加上条件：所考虑的空间都是 E^n 中的列紧体. 这个限制对于定理的证明有本质的意义，因为在证明中用了空间在 E^n 中的余集的性质，并且所有的推论都属于"Alexander-Понтрягин 对偶定律"一类的想法.

Л.С.Понтрягин[1]证明，对于所有的素数 p 及列紧体 X 与 Y，有下面的乘法定理

$$\triangle(X\times Y,\mathscr{G}_p)=\triangle(X,\mathscr{G}_p)+\triangle(Y,\mathscr{G}_p)$$

但是，对于 Урысон-Menger 维数论，这是不成立的.

关于此点，Урысон-Menger 维数并不是同调维数中最"好"的一个.

容易证明同调维数是"单调的"：若 X 为 Y 的下空间并且 $\triangle(X,\Gamma)$ 与 $\triangle(Y,\Gamma)$ 都有定义，则 $\triangle(X,\Gamma)\leqslant\triangle(Y,\Gamma)$.

60. 苏维埃拓扑学派的巨大成就之一是 Л.С.Понтрягин[4]所证明的亚历山大"对偶定律"的推广. 这个定律最初是对于 n - 维球中的多面体及系数域 $\mathscr{G}_2$





证明的.Л.С.Понтрягин 成功地推广这个定律到 n – 维球中任意闭集及作意贝蒂群的情形.后来,n – 维球又被成功地换为任意的“同调流形”.在没有叙述这个观念的定义以前,我们先得介绍“抽象同调流形”的观念作为其组合性质的基础.

看任意一个 A – 复合形 C.设 T 为它的任意一个间架.复合形 C 中所有的间架 S,满足条件

$$S \cap T = \Lambda, S \cup T \in C$$

的,其全体显然组成 C 的一个下复合形.这个下复合形倚赖于 C 及 T,称为 T 在 C 中的代表(Л.С.Понтрягин[8]).我们记作

$$P(T, C)$$

A – 复合形称为通的,假如它不能分为两个不相交的下复合形之和.若通 n – 维 A – 复合形 C 满足下列的条件,则称为抽象同调流形:若 r 为小于 n 的不负整数,T 为 r – 维间架,则有

$$B^s_\triangle(P(T, C), \mathscr{G})\begin{cases} = \{0\} & ,\text{对于 } s < n - r - 1 \\ \approx \mathscr{G} & ,\text{对于 } s = n - r - 1 \end{cases}$$

对于 n – 维抽象同调流形 C 必有:$B^n_\triangle(C, \mathscr{G}) \approx \mathscr{G}$或$\{0\}$.若抽象同调流形是属于第一种情形,则称为可定向的,若属于第二种情形,则称为不可定向的.对于所有 n – 维抽象同调流形必有:$B^n_\triangle(C, \mathscr{G}_2) \approx \mathscr{G}_2$.

现在设 X 为任意一多面体.由定义,它同胚于某一直多面体 Y.Y 是某一 E – 复合形 K 的所有的单纯形之和,而 K 又决定了一个 A – 复合形 C——由属于 K 的单纯形的顶点间架所组成.这样从已知多面体 X 来适出 A – 复合形 C 的做法并不是唯一的.不过,所造出来的 A – 复合形是 n – 维抽象同调流形这一件事却与构造过程中的任意选择无关,而由作为一个拓扑空间的多面体 X 所完全决定.这使得有可能来定义 n – 维同调流形作为那一种多面体,在上述的构成之下所得的是 n – 维抽象同调流形.显然,这样定义的 n 维同调流形的观念是拓扑不变的:所有同胚于 n – 维同调流形的空间必为 n – 维同调流形.以下我们提到 n – 维流形时都是指的 n – 维同调流形,提到 n – 维抽象流形时都是指的 n – 维抽象同调流形.

将抽象流形分类为可定向的及不可定向的也相应地产生了将 n – 维流形拓扑不变地分为可定向的 n – 维流形及不可定向的 n – 维流形两类.

61.为了陈述 Alexander-Понтрягин 对偶定律,我们适当地引入“无穷 A – 复合形”“无穷 B – 复合形”及“无穷多面体”等观念.

我们说间架所组成的集 C 是一个无穷 A – 复合形,假如它是一个备集,并且 C 中任意一个间架只包含于有穷个 C 中间架之内(“局部有穷”的条件).我们说由 E^r 中单纯形所组成的集 K 是一个 E^r 中的无穷 E – 复合形,假如它满足

E－复合形定义中的条件(1)与(2),并且 K 中单纯形的每一点必在 E^r 中有一近傍只同有穷个 K 中单纯形相交("局部有穷"的条件).我们说 E^r 中的某下集是一个无穷直多面体,假如它是 E^r 中某无穷 E－复合形的所有单纯形之和.我们说某一空间是一个无穷多面体,假如它同胚于无穷直多面体.无穷多面体的下开集也是无穷开多面体(龙格定理).

"无穷 $A(E)$－复合形"与"无穷多面体"的观念是"$A(E)$－复合形"与"多面体"观念的推广.许多关于复合形与多面体的观念和结果,自然地推广及于无穷复合形与无穷多面体,n－维 A－复合形的定义可以逐字地换为 n－维无穷 A－复合形.对于无穷 A－复合形 C,我们定义它的贝蒂群时将只考虑 C 的有穷 $<r,\gamma>$－链,即函数 f,其变数为 C 中 $r+1$ 个顶点,其值是 Γ 中的元,f 除了满足 $<r,\Gamma>$－链定义中的条件(1)与(2)以外,还得要满足条件(3):$f(x_0,\cdots,x_r)=0$ 仅对于有穷多组变数不成立.此处 Γ 可以为任意的陨群.算子△作用于这种链——如作用于有穷 A－复合形的链:我们得到 Γ 中元的无穷和,其中只有有穷多个异于零;这种和,我们也把它看做有穷和,无穷个零元略去不考虑.这样,我们就得到群 $L^r(C,\Gamma)$(C 中所有的有穷 $<r,\Gamma>$－链),$Z^r_\triangle(C,\Gamma)$(C 中关于 Γ 的 r－维有穷△－闭链,即 C 中之 $<r,\Gamma>$－链 f 而满足 $\triangle f=0$ 者),$H^r_\triangle(C,\Gamma)$(C 中关于 Γ 的 r－维△－闭链而为 C 中某有穷的 $<r+1,\Gamma>$－链的边缘的)以及 $B^r_\triangle(C,\Gamma)=Z^r_\triangle(C,\Gamma)\setminus H^r_\triangle(C,\Gamma)$.群 $B^r_\triangle(C,\Gamma)$ 称为无穷 A－复合形 C 关于系数域 Γ 的 r－维贝蒂△－群.

若 X 为无穷多面体,则它同胚于某一无穷直多面体 Y.Y 为 E^r 中某一无穷 E－复合形 K 的所有单纯形之和.属于 K 的单纯形,它们的顶点所成的间架,形成了一个 A－复合形 C.虽然这个复合形 C 并不是由多面体 X 唯一决定的,它的贝蒂群 $B^r_\triangle(C,\Gamma)$,除开一个同构,只倚赖于 X,而是 X 的拓扑不变量.这个群我们称它为多面体 X 关于系数域 Γ 的贝蒂群,记作 $b^r(X,\Gamma)$,$b^r(X,\Gamma)\approx B^n_\triangle(C,\Gamma)$①.

在 n－维抽象同调流形的定义中把 A－复合形换为无穷 A－复合形就得到无穷 n－维抽象同调流形的定义.以这个定义为基础,又可引到无穷 n－维同调流形的观念②,无穷 n－维流形 X 也分为可定向的与不可定向的两类,视是否 $B^n_\triangle(X,\mathscr{P})\approx\mathscr{P}$ 而定.无穷 n－维可定向流形的不变空间下集也是无穷 n－维可定向流形.

对于可定向无穷 n－维流形 X 有以下的庞加莱对偶定律

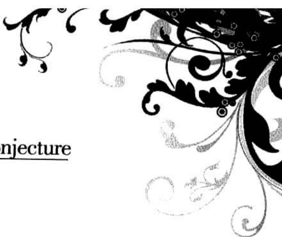
$$B^n_\triangledown(X,\Gamma)\approx b^{n-r}(X,\Gamma),\quad 0<r<n$$

---

① 群 $b^r(X,\Gamma)$ 不可与群 $B^r_\triangle(X,\Gamma)$ 混淆.

② 以下凡遇到与这个观念有关之处,形容词"同调的"将略去不写.





其中,Γ 为任意的陨群.

由此以及由 Колмогоров-Alexander 对偶定律,就得以 Alexander-Понтрягин 对偶定律:

若无穷 n-维可定向流形 X 关于系数域 Γ 在维数 r 与 $r+1$ 是单通的,则对于它的任意闭下集 Y 有

$$B^n_\triangle(Y,\Gamma)\mid b^{n-r-1}(X\setminus Y,G) \quad ⑧$$

其中,Γ 是紧群,G 是 Γ 的指标群,$0\leqslant r<n-1$.

Alexander-Понтрягин 对偶定理像这里所写的这样形式属于 П.С.Александров[17]①在对偶定理的这个形式中,Y 为闭集的假设是本质的:没有这个假设,一般而言,⑧中所出现的贝蒂群是未定义的.不过后来 Г.С.Чогошвили[1,2]及 П.С.Александров[20,21]相当地推广了这个定律,在某种程序上推广到 Y 不是闭集的情形.

62.空间与变换的同调性质不仅本身是有趣味的,并且与许多其他有趣的性质相关联.我们已经看到了同调性质与维数的关系.利用“形变”的观念来定义的“同伦”性质,与同调性质有重要联系.

设 f 与 g 为自空间 X 到空间 Y 的连续变换.我们说 h 为在空间 Y 中自 f 到 g 的形变,假如 h 是那样一个自空间 $X\times I$ 到空间 Y 的变换,满足 $h<x,0>=f(x)$,$h<x,1>=g(x)$对于 $x\in X$②.我们说在 Y 中 f 同伦于 g,假如有在空间 Y 中自 f 到 g 的形变存在.这样定义的在 Y 中的同伦关系显然是反射的对称的,以及可递的.因此,它就把自 X 到 Y 的连续变换分成了类:两个自 X 到 Y 的连续变换属于同一个类,当且仅当它们是同伦的.这些自 X 到 Y 的连续变换类,称为自 X 到 Y 的同伦类.所谓自 X 到 Y 的连续变换的同伦不变量,是指的那一种倚赖于自 X 到 Y 的连续变换的函数,它们对于任意两个互为同伦的变换的值是不变的.

在前一个时期,拓扑学中最出乎意料的发现之一就是 H.Hopf 的一件工作,证明自三维球 S_3 到二维球 S_2 的同伦类有无穷多个.为此,H.Hopf 对于自 S_3 到 S_2 的变换定义了一种数值的同伦不变量,并证明这不变量可以取任意整数值.Л.С.Понтрягин[7]完成了自 S_3 到 S_2 变换的同伦分类,证明了 Hopf 不变量是唯一的同伦不变量:

两个自 S_3 到 S_2 的连续变换为同伦的充要条件是它们具有相同的 Hopf 不变量.

Л.С.Понтрягин[5,6]也考虑了自多面体到球的连续变换同伦分类的一般问

---

① 在结论之中,$r=n-1$ 的情形是除外的.若流形 X 为 $n-1$ 维单通的,则对偶定律⑧对于维数为 $n-1$ 时也成立.

② 空间 $X\times I$ 的点看做有序的“点对”$<x,t>$,其中 $x\in X$,$t\in I$.

题.虽然,整个来说,问题尚未解决,他对于许多情形得到有趣的结果.在Л.С.Понтрягин所有这些研究之中,同调论担任了一个主要角色.

设 A 为空间 X 的下集.A 称为在空间 X 中可缩为一点,假如自 A 到 X 的全同变换在 X 中同伦于一个把 A 变换为 X 中一点的变换.最小的数目 n,使得 A 可以表示为 n 个在 X 中可缩为一点的集之和的,称为 A 在 X 中的畴."畴"的观念是 Л.А.Люстерник 与 Л.Г.Шнирелвман[1]所引入的,作为他们所研究的变分学中的拓扑方法等工作的基本工具①,那些方法也本质地用到了同调论.利用这些工具得到了变分学中许多有趣的新结果:只消列举著名的Л.А.Люстерник与 Л.Г.Шнирельман[2,3]定理关于在亏格为零的曲面上有三条闭测地线的存在.这里我们有着光辉的例子表明拓扑学其他数学分支间富有成果的相互影响.

设 X 为 E^r 中一个列紧体,U 为 E^r 中开集,Y 为 E^r 中的一个多面体.所谓 X 与 Y 在 U 中的分离是指自 $X\cup Y$ 到 E^r 的全同变换的一个形变 h,而满足

$$h(Y\times I)\subset U$$
$$h<x,1>=h<y,1>\quad ,x,y\in Y$$
$$h<x,t>=x\quad ,x\in X\setminus U,t\in I$$
$$h<x,t>\neq h<y,t>\quad ,x\in X,y\in Y,t\in I$$

我们说,列紧体 X 在 E^r 中产生 n - 维同伦障碍,在它的点 x 假如 x 在 E^r 中有一近傍 U 满足下列条件:在点 x 的每一个近傍之内都包含有 $(r-n-1)$ - 维多面体,与 X 不相交,且不存在有它与 X 在 U 中的分离.П.С.Александров[8]证明下面的定理:

E^r 中的列紧体在 Урысон-Menger 意义之下为 n - 维的充要条件是,它至少在它的一点产生 n - 伦同调障碍,而在任何点不会产生更高维的同伦障碍.

这个美妙的定理,借一种和列紧体在 E^r 中所处的位置有关的同伦术语予 E^r 中 n - 维列紧体以特征化,它的证明也用到了同调论,牵涉与 Alexander-Понтрягин 对偶定律属于同一个圈子的思想.

63.苏维埃拓扑学者的许多工作致力于所谓"开"变换.自空间 X 到空间 Y 的变换 f 称为是开的,假如 X 中任何开集的 f - 象,在 Y 中是开集.

众所周知,连续变换可以把维数增高,就是说,存在有实例说明连续变换可以把空间变到维数较高的空间上去.第一个这样的例是著名的皮亚诺连续变换自线段到方块上.自然地就产生这样的问题:开变换是否可以增高维数?换言之,是否存在开变换自某空间 X 到某空间 Y 上去,而 $\dim X<\dim Y$?А.Н.Колмогоров[6]给了这个问题肯定的答案,借助于谱论及 Понтрягин 列紧

① 这个问题的文献可参照 Л.А.Люстерник 的论文.





体,造出开变换把某一个 1 – 维列紧体变换到一个 2 – 维列紧体上.后来 Я.М. Каждан[1]造出开变换把某一 1 – 维列紧体变到方块上.

另一方面,П.С.Александров 证明了下面的定理:

若存在有可数重开变换把列紧体 X 变到列紧体 Y 上,则 $\dim X = \dim Y$.

我们说,自集 X 到集 Y 上的变换 f 为可数重,假若 Y 中每一点的 f – 反象的势都不超过 $\aleph_0$.

П.С.Александров 提出了下列未解决的问题:是否有开变换将 I^n 变到 I^m 上,$n < m$?

虽然这篇概览是很不完全的(由于篇幅的限制,许多为苏维埃拓扑学者所考虑过的重要问题都未曾提到),作者深信读者已经可以获得一种印象,苏维埃学者在近代拓扑学的每一分支都有着何等巨大的贡献,不仅在这门科学本体,而且在与其他数学各分支交错的地方.

参考文献

Александров,А.Д.

[1] 在抽象空间的加性集函(英文)[J].Матєм.сб.,1940,8(50):307-348.

[2] 在抽象空间的加性集函(英文)[J].Матєм.сб.,1941,9(51):563-628.

[3] 豪斯道夫空间之 H 闭推广(俄文)[J].ДАН,1942,37:138-141.

[4] 在抽象空间的加性集函(英文)[J].Маяєм.сб.,1943,13(55):169-238.

Александров,П.С.

[1] 紧拓扑空间的构造(德文)[J].Math. Ann.,1924,92:267-274.

[2] 局部列紧拓扑空间的计量化(德文)[J].Math.Ann.,1924,92:294-301.

[3] n – 维集合论的拓扑学之基础(德文)[J].Math.Ann.,1925,94:296-308.

[4] 论嘉当流形及 Phragmèn-布劳威尔定理的推广(法文)[J].C.R.Acad.Sci.,1926,183:722-724.

[5] 列紧空间的连续变换(德文)[J].Math.Ann.,1927,96:555-571.

[6] 论广义的维数观念及其与初等几何直观的关系(德文)[J].Math.Ann.,1928,98:617-636.

[7] 关于任意维数闭集的形状与位置的研究(德文)[J].Ann. of Math., 1929,30:101-187.

[8] 维数论.对于闭集几何学的贡献(德文)[J].Math. Ann.1932,106:161-238.

[9] 论拓扑空间叙列(法文)[J].C. R. Acad. Sci., 1935,200:1649-1651.

[10] 拓扑空间论(俄文)[J].ДАН,1936,2:51-54.

[11] 陨空间(德文)[J].Матем.сб.,1937,2(44):501-520.

[12] 论拓扑空间的紧推广(俄文)[J].Матем.сб.,1939,5(47):403-424.

[13] 论紧空间的维数(俄文)[J].ДАН,1940,26:627-630.

[14] 广义同调论(俄文)[J].М.,Учен зал,ун-та.1940,45:3-60.

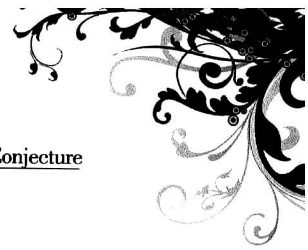
[15] 紧空间维数论的加法定理(俄文)[J].Тбилиси,Сообщ.Гр.фил.АН, 1941,2:1-6.

[16] 广义组合拓扑学(英文)[J].Trans. Amer. Math. Soc., 1941,49:41-105.

[17] 论复合形与闭集的位置的同调性质(俄文)[J].ИАН,сер.матем.,1942,6:227-282.

[18] 论复合形与闭集的位置的同调性质(英文)[J].Trans. Amer. Math. Soc., 1943,54:286-339.
[19] 论拓扑学中的空间观念(俄文)[J].Успехи матем.наук,1947,2:1(17):5-57.
[20] n-维空间中非闭集的广义对偶定律(俄文)[J].ДАН,1947,57:107-110.
[21] n-维空间中非闭集对偶性的基本定理(俄文)[J].Матем.сб.,1947,21(69):161-231.
Александров,П.С.与 Урысон,П.С.
[1] L空间是D空间的一个充要条件(法文)[J].C. R. Acad. Sci., 1923,177:1274-1276.
[2] 拓扑空间论(德文)[J].Math. Ann., 1924,92:258-266.
[3] 紧拓扑空间论(法文)[J].Amsterdam. Verh. Kon. Akad. Wet., 1929,14(1):1-96.
[4] 论紧空间(法文)[J].Bull. Acad. Polonaise. (A) 1929,5-8.
Александров,П.С.,Понтрягин,Л.С.与 Hopf, H.
[1] 论布劳威尔维数观念(德文)[J].Comp. Math., 1937,4:239-255.
Вайнберг,Н.М.
[1] 论拓扑空间的常闭性(俄文)[J].ДАН,1941,31:523-524.
Веденисов,Н.Б.
[1] 短论拓扑空间的连续函(俄文)[J].М.,Учён.зап.пед.ин-та,сер.физ.матем.,1938,2:47-54.
Каждан,Я.М.
[1] 自一维局部通连续统到方块上闭变换的实例(俄文)[J].ДАН,1947,56:339-342.
Колмогоров,А.Н.
[1] 论组合拓扑学中的对偶结构(德文)[J].Матем.сб.,1936,1(43):97-102.
[2] 局部紧空间的贝蒂群(法文)[J].C.R.Acad. Sci., 1936,202:1144-1147.
[3] 局部紧空间贝蒂群的性质(法文)[J].C.R.Acad. Sci., 1936,202:1325-1327.
[4] 计量空间的贝蒂群(法文)[J].C.R.Acad. Sci.,1936,202:1558-1560.
[5] 相对闭链.亚历山大对偶定理(法文)[J].C.R.Acad. Sci.,1936,202:1641-1643.
[6] 论开变换(德文)[J].Ann. of Math., 1937,38:36-38.
Курош,А.Г.
[1] 紧拓扑空间的组合结构(德文)[J].Comp. Math., 1935,2:471-476.
[2] 有穷集所形成的半序系理论(俄文)[J].Матем.сб., 1939,5(47):345-347.
Люстерник,Л.А.
[1] 函数空间的拓扑学及大范围变分学(俄文)[J].Труды матем. ин-та им.Стеклова,1947,19:1-96.
Люстерник,Л.А.与 Шнирелвман,Л.Г.
[1] 论分析中的拓扑原理(法文)[J].C. R. Acad Sci., 1929,188:295-298.
[2] 亏格为零的曲面上存在三条闭测地线(法文)[J].C. R. Acad. Sci., 1929,188:534-537.
[3] 论亏格为零的曲面上必存在三条闭测地线的问题(法文)[J].C. R. Acad. Sci., 1929,189:269-271.
Марков,А.А.
[1] 关于阿贝尔集的某些定理(俄文)[J].ДАН, 1936,1:299-302.





[2] 论积分不变量的存在(俄文)[J].ДАН,1937,17.
[3] 论均值及外密度(英文)[J].Матем.сб.,1938,4(46):165-191.
[4] 论自由拓扑群(俄文)[J].ИАН,сер.матем.,1945,9:3-64.
Понтрягин,Л.С.
[1] 论维数论中的一个基本假设(法文)[J].C. R. Acad. Sci., 1930,190:1105-1107.
[2] 对于闭集的广义对偶定律(德文)[J].I nt. Mathematikerkongress Zürich, 1932,2:195-197.
[3] 可交换拓扑群论(英文)[J].Ann. of Math. 1934,35:361-388.
[4] 对于闭集对偶性的一般拓扑定理(英文)[J].Ann. of Math., 1934,35:904-914.
[5] 自复合形到球上连续变换的分类,Ⅰ.(俄文)[J].ДАН,1938,19:147-150.
[6] 自复合形到球上连续变换的分类,Ⅱ.(俄文)[J].ДАН,1938,19:361-364.
[7] 自三维复合形到二维球变换的分类(英文)[J].Матем.сб.,1941,9(51):331-364.
[8] 拓扑的对偶定理(俄文)[J].Успехи матем. наук,1947,2:2(18):21-44.
Понтрягин,Л.С.与 Шнирельман,Л.Г.
[1] 论维数的一个计量性质(法文)[J].Ann. of Math., 1932,33:152-162.
Тихонов,А.Н.
[1] 论 Урысон 的一个计量化定理(德文)[J].Math. Ann., 1925,95:139-142.
[2] 论空间的拓扑推广(德文)[J].Math. Ann., 1929,102:544-561.
[3] 论自紧空间到欧几里得空间的变换(德文)[J].Math. Ann., 1935,111:760-761.
[4] 一个不动点定理(德文)[J].Math. Ann., 1935,111:767-776.
[5] 论万有拓扑空间(俄文)[J].ДАН.1936,3:49-52.
[6] 论反值问题的稳定性(俄文)[J].ДАН,1945,39:195-198.
Тумаркин,Л.А.
[1] 论广义维数论(德文)[J].Proc. Amsterd. Akad., 1925,28:10.
[2] 对广义维数论的贡献(德文)[J].Матем.сб.,1926,33:57-86.
[3] 论闭集的维数结构(法文)[J].C. R. Acad. Sci., 1928,186:420-422.
[4] 论非闭集的维数(德文)[J].Math. Ann., 1928,98:637-656.
Урысон,П.С.
[1] 论列紧拓扑空间的计量化(德文)[J].Math Ann., 1924,92:275-293.
[2] 看做计量空间反象的希尔伯特空间(德文)[J].Math Ann., 1924,92:302-305.
[3] Canter 流形详论(法文)[J].Fund. Math., 1925,7:29-137.
[4] 论通集的势(德文)[J].Math Ann., 1925,94:262-295.
[5] 计量化问题(德文)[J].Math Ann., 1925,94:309-315.
[6] 嘉当流形详论(法文)[J].Fund. Math., 1926,8:225-351.
Фомин,С.В.
[1] 拓扑空间推广论(俄文)[J].Матем.сб., 1940,8(50):285-294.
Чогошвили,Г.С.
[1] 收缩核的对偶定律(俄文)[J].ДАН, 1946,51:87-90.
[2] 无穷多面体的对偶定理(法文)[J].C. R. Acad. Sci., 1945,221:15-17.
[3] 拓扑空间中的对偶关系(俄文)[J].М., Диссертадия, Матем. ин-т им. Стеклова,1947.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

Шанин, Н. А.
[1] 论拓扑空间的特殊推广(俄文)[J]. Матем. сб., 1940, 8(50): 285-294.
[2] 论拓扑空间的分离性(俄文)[J]. ДАН, 1943, 38: 118-122.
[3] 拓扑空间的紧推广论(俄文)[J]. ДАН, 1943, 38: 166-169.
[4] 一般集论的一个定理(俄文)[J]. ДАН, 1946, 53: 403-464.
[5] 论拓扑空间积的开下集的相交(俄文)[J]. ДАН, 1946, 53: 503-506.
[6] 论拓扑空间积(俄文)[J]. ДАН. 1946, 53: 595-598.
[7] 论二进紧体(俄文)[J]. ДАН, 1946, 53: 785-788.



下　编

面向大众的拓扑学描述



面向大众的拓扑学描述

第十七章

1 塞吉·朗关于拓扑学的演讲①

这是我第三次来这里，巴黎科学博物馆，和你们一起做数学.第一次是布雷特先生邀请了我，效果不错，所以我又来了.

我发现在座有不少人上次也来过，有多少人去年也来过？

［约 50 人举手］

很好.我看到安图望坐在那里，他已经来过两次了，他是很忠实的听众.去年来过的听众可能记得，演讲会开始前我在布雷特的办公室翻了翻一些高中课本，看了这些令人作呕的课本我很不舒服，过了 20 分钟才稍好一点.不知道你们是否注意到，今天的演讲会前，你们听到了一首诗琴曲的录音，那是我最喜欢的乐曲.布雷特放这些音乐可以使我平静下来.［笑声］

两年前，我们讨论了素数的某些性质，去年，我又讲了丢番图方程，已经印在这个预印本上.我问过听众，数学对他们意味着什么？一位女士回答说："是跟数字打交道."好，这个回答是牛头不对马嘴，因为这完全不是做数学的意思.我希望能告诉你们数学是什么，数学中的大问题是什么，以及为什么人们会为之振奋.

实际上，在头两个演讲中，我讲的事或多或少与代数有关.

① 原载：［法］塞吉·朗.做数学之美妙［M］.李德琅，译.郑秋成，校.成都：四川大学出版社，2001.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

特别是去年，我写下了某些公式，而有6位听众立即退场，因为公式……是的，人们不大喜欢公式.但有时候，公式是必要的.而且，我怀疑，不用任何公式，不涉及任何代数，不涉及任何数字，我们还能做什么事呢？我想，这只能是某种几何的东西，空间中的东西，与几何物体相关的问题.这不是我自己平常做的数学.就个人而言，我偏爱代数及数论.因此当我离开巴黎去德国波恩时，我考虑了这个问题，试图为今年的演讲物色一个话题.过去的20或25年中，我每年都去波恩.希策布鲁赫（Hirzebruch）主持一个会议，参加者主要感兴趣于几何课题.我与他们的交谈使我意识到可以选这样一个课题，关于大约一年前发现的某些研究成果.

在波恩真是非常愉快.数学家们在欢快的气氛中举行会议，一边喝莱茵酒、吃草莓蛋糕——草莓正是季节，一边做数学，而且还在莱茵河上乘船漫游.

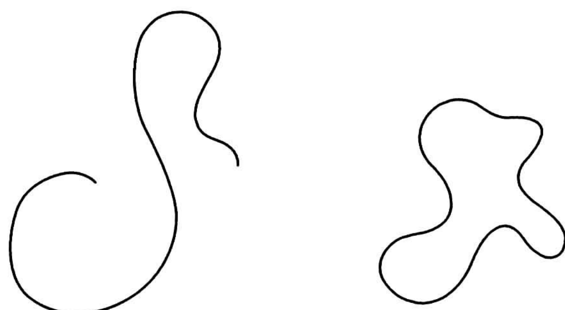
当然还是有时间做数学，在那里我学到了今天准备讲的这个课题，一位叫瑟斯顿的伙计的某些新发现.我是从沃特诺伊曼（Walter Neumann）那里学到这些的.我们在黑板前花了三个小时，他给我讲瑟斯顿做了什么.今天，我把这些传达给你们.原先我真不知道我也可以这样做，选择别的话题，因为这并不容易.首先，它必须是真正的数学，真正的数学家的工作.其次，它必须是能够向周末演讲会的听众讲清楚的.还有，像所有审美情况一样，有些人会喜欢，有些人则相反.也许它不灵.这是一个个人情趣的问题，你们对任何话题都可以有自己的反应.

好了，我今天要讲的，是几何对象的分类.这立即会引起我们的兴趣.我们生活在至少三维的空间中，就像这样.但是，你们已经知道还可能有多于三维的空间.所以我们想描绘我们生活中的空间，想知道它究竟像什么样子.局部地说，例如，这间大厅，它就是三维空间.这样一个模式，用来考察这间大厅是足够了.但我们知道，如果你看得更远些，就不同了.我们知道，欧几里得模式其实是不对的.它在某些限制下是对的，但不适用于其他一些情形.那么，物理学家们怎么办呢？他们试图找出，哪些模式是适用的.但是数学家，纯粹数学家，不在乎这些模式是否适用.他们构造出好的模式，当然是几何模式，只要感到美妙就行.他不在乎这个模式是否可以用以描绘宇宙.

我们可以做一维的几何模式，也可以做二维的、三维的，以及更高维的，像四维、五维的等等.我原想今天谈谈高维的东西，但我很快就意识到，在一个半小时内我是办不到的，那会有太多的东西需要准备.所以，今天，我仅仅限于谈一维、二维和三维的.那么，一维的东西就像这样，是曲线.

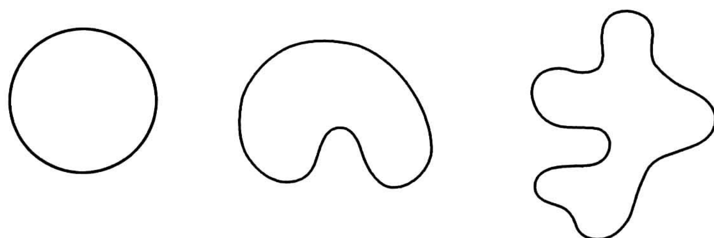






一女士:那直线呢?

塞吉·朗:直线是曲线的特别情形.现在如果我画一个圆,以及其他曲线,就像这样,它们看起来彼此相像.



我们是否应该认为它们等价呢?它们有什么共通的性质呢?

一听众:它们是闭合的.

塞吉·朗:是的,它们是闭合的,它们转了一圈.如果我取一条线段,像这样:

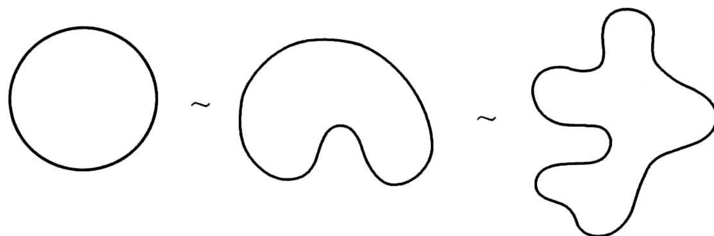
———————

那就不能转一圈,它是一条线段.

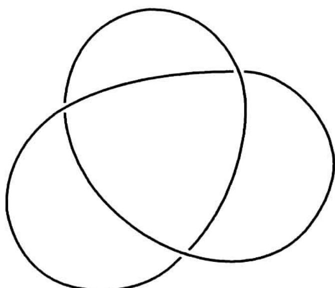
但上面三条曲线是闭合的.对于许多性质来说,我们不想区分这三条曲线.所以,我们说这些曲线是等价的.

一般说来,等价的意思是什么呢?啊,我不想给一个正式的定义,但我可以非正式地说,我们假定它们都是橡皮做成的,我们考察橡皮几何.我们说两个东西等价,如果可以从一个方向拉一拉,另一个方向压一压,如果它们都是橡皮的,那可以把一个变成另一个.这就给了我一个等价的概念.

如果这些曲线是一条橡皮绳,显然我可以把其中一条变成另一条,或者我可以把它变成一个圆.所以这些曲线是等价的.我用符号"~"来表示等价.我可以这样记:



我还可以这样画一条线:



这是闭合的还是开的?

听众:是闭合的.

塞吉·朗:它等价于原来那几条吗,假定它是橡皮绳?谁说它等价于其他几条?[一些人举手]

哪些人说它不?[另一些人举手]

哪些人保持审慎的沉默?[笑]例如,你.[塞吉·朗指向第三排的一位女士]

女士:它不等价.这里有一个结.

塞吉·朗:是的,有一个结.当我说那几条曲线等价时,我可以在平面上把它们变来变去.我的意思是,如果它们是橡皮绳,我可以完全在平面上完成变形.但这个结,就在那里,它只存在于三维空间中.你的直觉是对的:我不能在三维空间中把它变成一个圆.从某种意义来说,结不同于圆,也不同于另两条曲线.但是,你是否相信在有一种情况下,我可以把一个结变成一个圆?安图望,你想说什么?

安图望:[从录音带上听不清他的回答]

一女士:有时你可以打两个相反的结,然后同时解开.

塞吉·朗:现在,这个结是在三维空间.但是没有理由把我们自己限于三维空间.事实上,在四维空间中我们可以把一个结变形为一个圆.也可以证明,在三维空间中这是办不到的.虽然,在三维空间中,我们可以依赖直觉,而要在高维空间中证明某件事,我们必须把事情用更严格的方式写出来,因为我们的直觉变得有些无能为力.我还想让你们明白,事情不是那么简单.

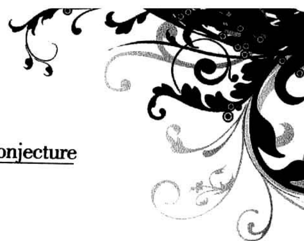
我们看到,可以提出两个不同的问题:

在三维空间中可以把结变成一个圆吗?

可以抽象地变形吗?在高维空间中可以变形吗?

答案是不同的,这取决于我们把这个结放在什么空间.起初,当我定义等价的概念时,我没有说你可以在哪里进行变形.现在我说我允许在任意的,高于二维及三维的高维空间中进行变形.于是,曲线的维数总是1,应该看做它完全不同于我们考察这个圆所在的空间的维数.现在,我想对变形再说点别的什么.取



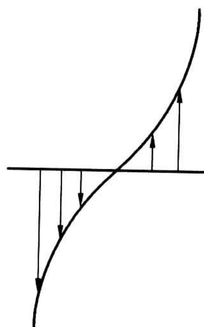


一条不是圆的曲线,例如说,一条线段,就像这样,包含或者不包含端点.



包含端点　　　　不含端点

如果包含端点,我说这条线段是闭的.如果不包含端点,那我说这条线段是开的.假定线段是用橡皮做成,我把它变形,就像这样.[塞吉·朗边说边画]



我让右边的点往上动,而左边的点往下动.所以我拿这条橡皮绳,我往右走时,把它往上拉,越来越快.而我往左走时,把它往下拉,也是越来越快.于是我们看到,这条线段等价于一条曲线,它可以跑到任意远,就像人们常说的跑到无穷远了.

[有人举手]

塞吉·朗:请.

一女士:它在无穷远处会闭合起来.

塞吉·朗:不,无穷远不是一个点.取一直线,像这样:

这条直线是不会闭合的.

一男士:如果它是橡皮做的,可以把它闭合起来.[笑]它不是一条线段,但含有一条线段,你可以把它闭合起来成一个圆.

塞吉·朗:看!如果你闭合一条直线.或者你闭合一条线段,那你必须放一些点在端头.取一条线段包含端点.如果我把端点连在一起,我得到一个圆.

男士:但线段可以变形为一个圆.



塞吉·朗:不,不是.因为如果我把它变形为一个圆,而且把两个端点视为同

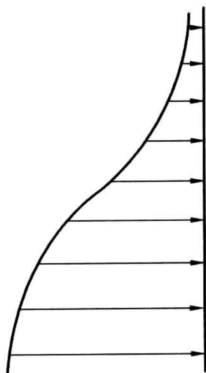
一个点，那我就干了某种为变形定义所不容的事．我使用变形一词是在不允许同一的意义下，即，不能够把两个点视为同一个点．如果两个点不同，那它们经过变形也互不相同．

男士：但如果你把它们并置一处——

塞吉·朗：不，不，我不想这样做．[笑声]这是一个定义的问题．我现在想用"变形"一词去指：如果两个点不同，经过变形它们也互不相同，行吗？

男士：行．

塞吉·朗：好．当然，也有其他的概念，那里允许同一．其实，待会儿我就要讨论这种概念，以及如何去使用它们．但是，现在，对变形，我不允许同一．我正是想让你看到这种特别的现象，可以把不含端点的线段变形为无穷长的一条绳，它本身就等价于无限长的直线．我可以把这条无限长的绳子与无穷直线的等价性画出来，就像

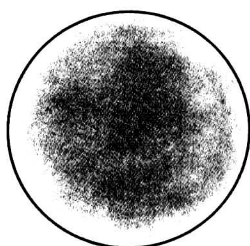


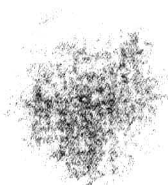
所以我可以把这条曲线拉直成一条直线．那就是我想要讨论的等价概念．

好了，我们刚才谈的是一维的事情．但是，即使是一维的情形，我们看到已经可以提出若干问题来．你也许觉得一切都是已知的，但事实并非如此．接下来，我们看二维的情况．这稍稍复杂一点．一维的东西称为曲线．二维的东西称为曲面．而且曲面可以含有边界，也可以不含边界．

作为曲面的一个例子，取一个圆盘，即圆的内部．如果我把圆及其内部放在一起来观察，我得到一张含有边界的曲面．圆就是圆盘的边界．所以只要去掉这个圆，我们可以把圆盘看成不含边界的曲面．如果你包含这个圆在内，也可以把它看成有边界的曲面．



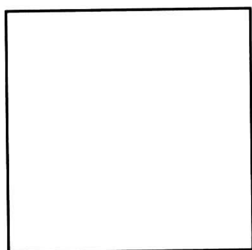




有边界　　　　无边界

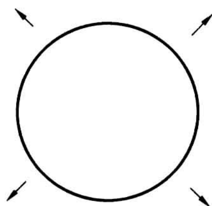
现在设这个圆盘是橡皮做的,那我就可以把它用其他方式来表示,例如,我可以取一个正方形的内部.它的边界就是正方形的周边.

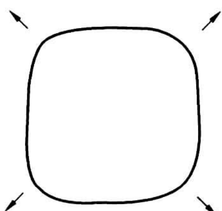
如果都是橡皮做的,它们等价吗?

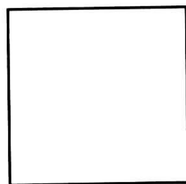


听众:是的.

塞吉·朗:对了,它们等价,可以取一个圆盘,把它拉开得到一个正方形,正方形的内部.







而圆盘的边界,就是这个圆,会变成正方形的边界.

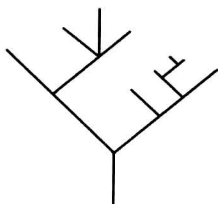
[有人举手]

塞吉·朗:请.

男士:但这里有某种不同之处,因为导数——

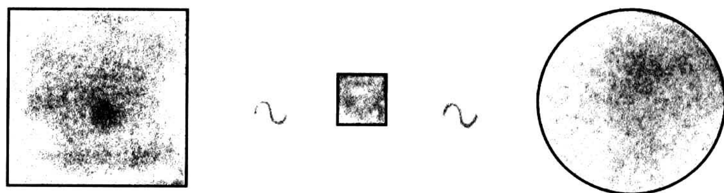
塞吉·朗:当然,你是指几个角.这位男士说这里有些区别,他完全正确.这里是有区别,但这不是橡皮几何的观点.可以定义其他类型的等价,在这新的定义下,圆盘和正方形这两个东西并不等价.因为,当我拉开圆盘并且做成一个角时,显然,这个角是不那么平滑的.从某种观点来看,你甚至可以说这个角是很不舒服的.[笑声]它不光滑,它不是一条好的曲线.从某种意义上说,它们是不

同的.还有关于角部的数学理论.你看,我们从某些相当简单的事情出发,已经列出了大量的问题,它们的发展就像一棵树:



我们爬上这棵树,发现两种甚至多种可能性可以继续下去.按照你考察的不同的等价关系,对同一个问题可以有不同的答案.但是现在,我只想考察橡皮等价关系.于是,圆盘与正方形是等价的.

当然,这不依赖于它们的大小.我可以把正方形画得大点或小点.只要它是橡皮做的,它都等价于圆盘.



如果我只取圆盘的内部,那我得到一个没有边界的曲面,或者没有边界的正方形.这有点像没有端点的线段.你还记得没有端点的线段吗?现在我来看正方形的内部,或者圆盘的内部,这是等价的,不考虑边界,同时来看平面,它从各个方向延伸到无限远.请想想正方形的内部是否等价于平面?



哪些人说是的?

一男士:平面是无限的.

塞吉·朗:是的,平面是无限的.

男士:这个正方形没有边界?

塞吉·朗:对,它没有边界.我把它去掉了,那正是我用虚线画的原因.

男士:那它也是无限的?

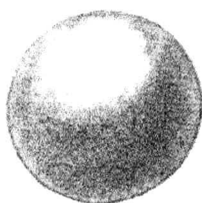
塞吉·朗:正如你所说,无限的.





男士:那它们等价.

塞吉·朗:很对.去掉了边界的正方形与平面等价.总结一下,每个去掉了端点的线段等价于无限长的直线,每个去掉了边界的正方形或者圆盘等价于整个平面.但是请留意:我取了一个去掉了边界的正方形,如果我愿意,我可以把边界加上去.但是,如果我取一个球面,像这样,一个球的表面:



它是一个曲面,但它没有边界,是吗? 如果我拉伸它,是否可以拉到使它的一部分拉到你想要多远就多远呢?

男士:你可以无限地吹胀这个气球.

塞吉·朗:请看,我不想把这个气球弄破.[笑声]这个东西必须保持等价.我吹胀这个气球,并且把它往里敲敲或往外拉拉,就像橡皮,但我不允许把它搞破.如果我拉长它,是不是可能像我们拉一条去掉端点的线段一样,把球面的一部分弄到你想多远就多远呢? 哪些人说是的?



[一些人举手]

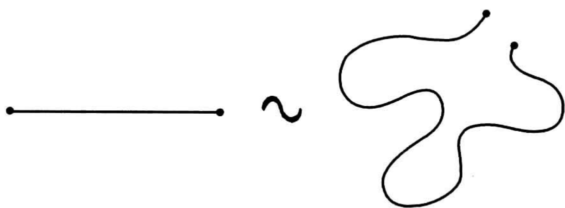
哪些人说不是? 其实,答案是"不".例如,如果我取一个带端点的线段,是否能把它拉长到等价于某种无限的东西呢?

一女士:应该是有界的.

塞吉·朗:对了.可以证明这是不可能的,带有端点的线段不能等价于无限的东西.

女士:你是否指端点是固定的?

塞吉·朗:啊,不! 它们不必被固定起来,你可以移动它们:例如,它可以等价于这种东西.



只要拉一拉，压一压，扯一扯就行．但问题是可不可能拉得越来越快，就像我们以前拉线段那样．问题出在如果我拉得越来越快，那端点就将无处可去．先前，线段上越接近终端的点或者跑得越来越高，或者越来越低．因而，如果在变形过程中把终端加上去，我只得把端点撕掉．但这是不允许的．

一男士：你可以把端点放到无穷远处．

塞吉·朗：不，我们必须在平面上干这事，平面上没有无穷远点，当然有任意远的点，但这并不是一回事．

男士：为什么禁止用它呢？

塞吉·朗：禁止用它是为了等价概念的定义．原则上并不是总禁止用它，并不是绝对禁止用它．对其他的某些应用而言，你可以把无穷远点添到平面上去，但不是我们今天所要做的．所以，你必须区分的是，有些东西有这样的性质，你可以经过变形把它的一部分送到任意远去；而有些东西不具有这样的性质，就不可能这样做．好，让我写一个定义：

> 说某物是紧致的，是指它包含它的边界（如果存在的话），而且此物的变形过程不能把它扩张到任意远处．换言之，此物的每个变形都是有界的．

归根到底是说球面是紧致的．当然，我们生活于其中的三维空间会跑到无穷远……［犹豫］起码按我们心中的质朴的模式，它会跑到无穷远去．但若你生活在一个球面上，而且你又非常小，当你看你的周围时，从各个方向看出去，你看到的很像一个平面……

［有人举手］

塞吉·朗：请．

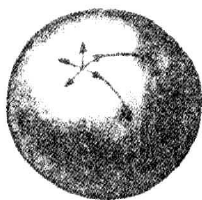
一大学生：而球面没有边界，你说"没有边界的紧致……"

塞吉·朗：啊！如果曲面没有边界，那它自然包含了它的边界．这个术语必须这样来理解：如果某物没有边界，那它不得不包含它的边界，因为根本就没有．［笑］你必须允许这种可能性，否则，一些简单数学命题都会很难办．

让我们回到住在球面上的人．也许他们只能看到一个平面，即使他有高倍望远镜在手；于是他们会很快得出一个结论，说他们住的空间是一个平面．如果



一千年以后,他们造出了更高级的望远镜,也许他们会发现某种弯曲,他们会看到,空间是弯曲的,并且开始提出一系列问题.



这正是哥伦布以前发生的事.人们设想什么都是平的,只有少数聪明人例外,但聪明人并不多.

听众:多新鲜![笑声]

塞吉·朗:好,我们看起来就是那样.我们还可以问,如果我们一直往前走,会发生什么事,我们会回到出发点,还是会无限地走下去?球面是紧致物体的例子.如果你从某点出发,沿着一个固定的方向一直往前走,那你最终会回到你的出发点.

你能不能给我另外一个这种曲面,紧致曲面的例子?

听众:立方体.

塞吉·朗:是的,立方体的表面,但它等价于球面.给我一个例子,要它不等价于球面.

一男士:环面.

塞吉·朗:什么?

男士:环面.

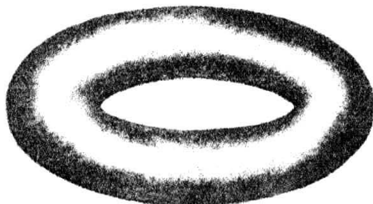
另一人:在球面上打个洞.

塞吉·朗:你是数学家吗?

男士:知道一点.

塞吉·朗:那已经太多了.我希望数学家不要介入,要不然就是欺骗.[笑声]当然数学家知道答案,但我这不是给他们作的演讲.[塞吉·朗把粉笔扔向那位男士.笑声]

所以,你开了一个洞,你找到了这个东西,它有一个洞在中间.



可以证明这个曲面不等价于球面,因为有这个洞.现在,你能不能给我一个曲面的例子,它不等价于球面或者环面?

一人:克莱因(Klein)瓶.

塞吉·朗:你们中间有些人知道得太多了.①

一小孩:金字塔.

塞吉·朗:不,它等价于球面.②

一女士:一个没有顶部的盒子.

塞吉·朗:是的,但它有边界.我想要一个没有边界的曲面.前面谈到的这个没有,球面也没有,我要一个紧致曲面.

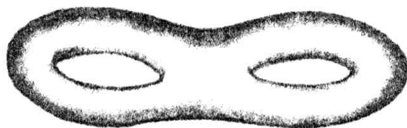
一大学生:你可以开两个洞,像眼镜架.

塞吉·朗:你说对了,那就是我希望你们说的.但是,你是不是学数学的?

学生:是的.

塞吉·朗:啊,不,不,不要这样.自然,如果你是数学家,你会说:开两个洞.但你不该参与这个游戏.[笑声]那就是为什么我不要你们介入.我想让人们自己想出来.

好,你说对了,我可以开两个洞.像这样:

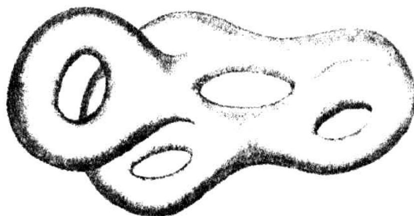


如果我还想要另一个例子,该怎么办?

一学生:环面上打个结.

一女士:你可以开更多的洞.

塞吉·朗:非常好.女士,第一次演讲你来了吗?两年前?你记不清了?我把你记得很清楚.总之,你可以开更多更多的洞.只要是有限多个洞.



① 我不想现在深入到这样的技术细节.

② 听众中有各方面人士,包括12岁的孩子、高中生、大学生、工程师和退休人员.后来我知道这小孩是12岁.她的老师要求她的同班同学们来听演讲并在事后写一篇印象.这女孩写道:当然,有时我多少有点糊涂了,例如当兰先生要大家举一个与球面不同的例子时,我说"金字塔",因为我知道兰先生想要一个类似的例子.此外,一切都好.

另一人写道:

如果我事先知道我的发言会被写在书里,我一定会多举几次手.



于是,有一个定理:

紧致无边界曲面,在等价的意义下,可以由洞的个数完全刻画.没有其他的紧致无边界曲面.

我还需要在定理的陈述中加上一个假设.我应该说:一个可定向的曲面.但我现在不想深入到这种问题去了.所以,忘掉我刚才所说的.如果我不说到它,就有些人会提出挑剔,比如某位数学家.[笑声]

[当塞吉·朗想在黑板上写出定理时,发现没地方写了,他不得不擦去黑板上的另一部分]

啊,没地方写了! 所以你们都应该写信给教育部部长,要他给科学博物馆的布雷特先生更多资金,以便他能够买更多的黑板,更大的黑板,放在一个更大的厅里,诸如此类……如果可能的话,非紧致资金更好.[笑声]你们在会后都去写信,而我来写这个定理:

> 可定向的(只是为使我安心),无边界的紧致曲面在等价的意义下由洞的个数刻画.

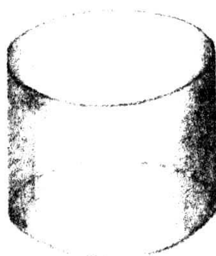
这就是曲面的一般模式.

现在我们来看有边界的曲面.

一男士:会不会只留下洞呢?

塞吉·朗:总会有曲面留下来的.我做这些都是为三维物体作准备,三维时事情更复杂.

好,我画一个有边界的曲面.刚才已经有人提到了圆柱面.



什么是这个圆柱面的边界?

一男士:一个圆.

塞吉·朗:对,有一个圆在顶上,还有一个圆在底部.圆柱面的边界由两个圆组成.

一女士:还有某些边.

塞吉·朗:不对,你让圆柱面转一周,你在旁边看它,你看不到有边.

女士:那它有两个边界.

塞吉·朗:是的,或者再说清楚一点,它有一个边界,但由两个圆组成.没人说过边界只能由一片组成,它也不必是连通的.

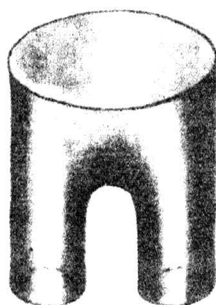
现在,我要画一个更好玩的.谁能告诉我怎样画一个曲面,其边界由多于两片组成?

一男士:人的脸.

塞吉·朗:是的,可作一个例子.

一女士:一个筛子.[笑声]

塞吉·朗:是的,非常好.让我再画一个.

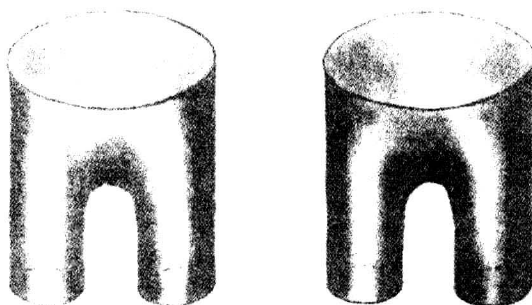


这是什么?

一男士:一个倒过来的花瓶.

其他听众:一条裤子.

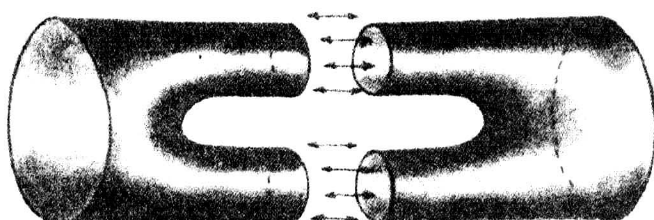
塞吉·朗:是的,一条裤子.它的边界包括顶上的一个圆和底部的两个圆.所以边界包括三片.现在我要做一件数学家们爱做的事.数学家喜欢把事情放在一起而做加法.假设你有两条裤子.



我能对它们做些什么呢? 如果我在它们上面各取一个圆,我可以把它们粘上.







我对另一只裤腿也做这种黏合.这样,我得到一个新玩意,可以称为两条裤子之和.

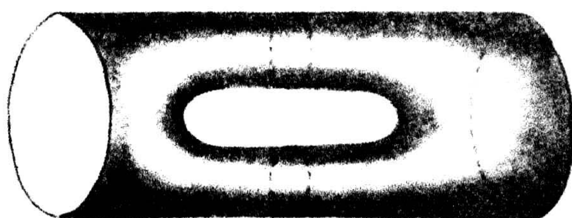
一男士:但你没有权利这样做,你是在把不同的东西视为同一.

塞吉·朗:现在我有这个权利.我在做加法,我把两条裤子缝到一起.[笑声]我有权做裁缝.我们到了可以谈同一的时候了.

女士:为什么刚才你还没权利搞同一,现在你就有权搞同一了?

塞吉·朗:你永远有权搞同一,把两个点放在一起,做你想做的.但是,为了什么目的? 为了定义等价,你无权这样做.我没有说,我搞了同一则我得到一种等价.我只是说我得到了和.这是不同的事情.对等价性,我不允许不同的点同一.对求和,我有权这样做.求和就是对边界的一些片搞同一.

所以,如果我把和画出来,我得到某种像这样的东西,有一个洞,仍然有边界,由两个圆组成.

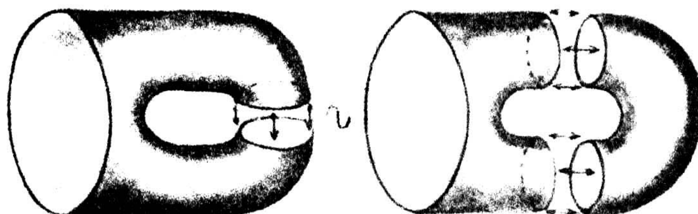


现在我可以再求和,消掉这两个圆.

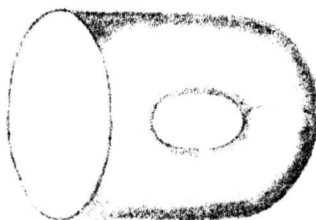
一男士:总而言之,你把两个不同东西的边界搞同一.

塞吉·朗:是的,两个不同的东西.我取两条裤子,你的和我的,然后我把它们缝在一起.[笑声]

现在,我也可以取一条裤子来求和,就是把两条裤腿上的两片边界搞同一.它也给我一个有洞的曲面.



如果我把这个曲面与圆柱面求和,我得到的还是这个曲面,它们是等价的.我得到一个洞,而且还有一个圆作为其边界.

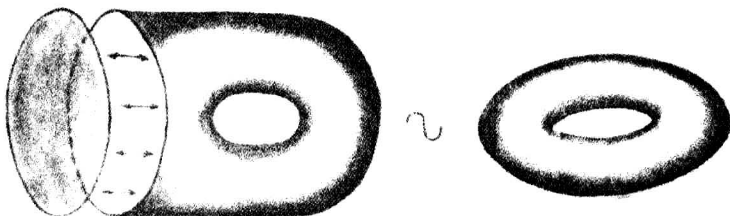


现在看这个圆,我想把它消去.我该怎样做?把它与什么东西求和就可以完全消去这个边界?

男士:半个球面?

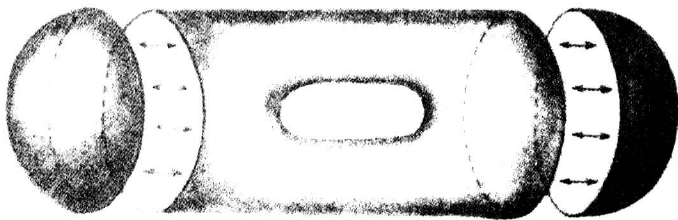
女士:一个盖子.

塞吉·朗:完全正确.一个盖子,它就是一个有边界的圆盘.



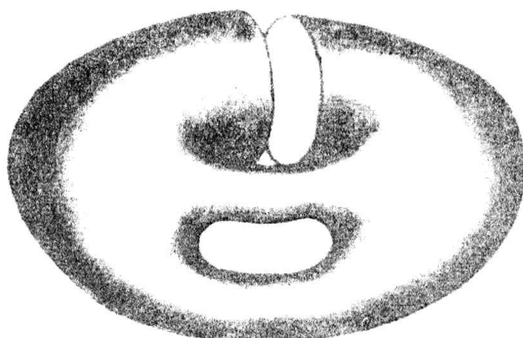
你把它们粘上,你得到某种没有边界的东西.这就是我想做给你们看的.如果我取一些有边界的曲面,其边界是圆,我作它们的和,我还可以把它们加若干次,最后得到一个有洞但无边界的曲面.

再来看裤子.我可以消去边界.先把两条裤子的裤腿缝到一起,然后在两端各放一个盖子.我得到一个环面.

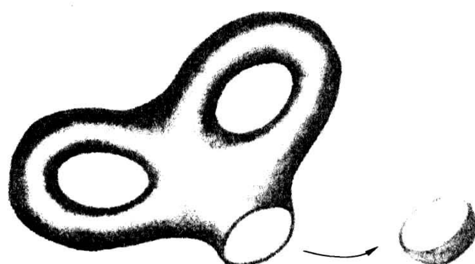


一学生:但你也可以把它们沿背带缝到一起.

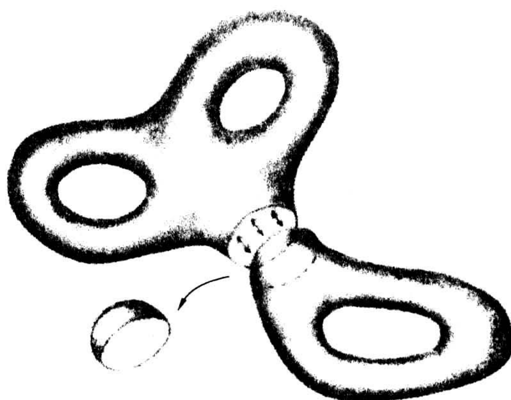
塞吉·朗:是的,可以的,不过你会得到更多的洞.



数学家们喜欢做这种事,这是一件使他们高兴的事情.[笑声]如果你也乐于把裤子缝到一起,那按照定义,这是在做拓扑(topology),而你就是一位拓扑学家.即使没有边界,你也可以造出一个边界来,以定义另一种加法.直至目前,我们只做了缝合,但也可以做外科手术.取一个曲面,像这样,很光滑,没有边界.于是我切下一个圆盘.

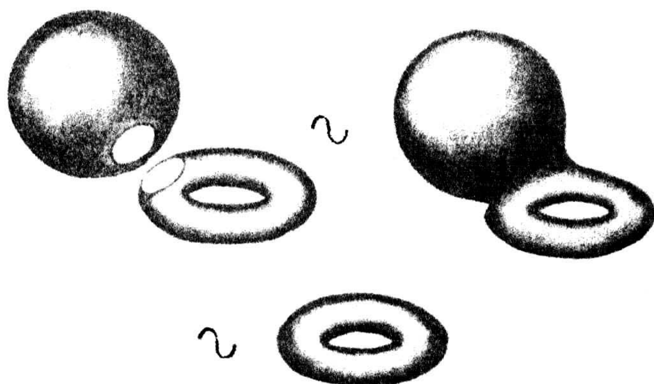


这就造出一个边界来,它是一个圆,原来没有的.我想对另一曲面干同样的事,造出另一个圆.现在我处在与先前同样的地位,我有两个有边界的曲面,我可以沿着这两个圆作它们的和.



用这种办法,我可以定义无边界曲面的和.如果我用任何曲面与球面求这

种和,我得到同一个曲面,当然是在等价的意义下得到的.人们称球面是这种求和法的零元素.



另一方面,如果我有一个带两个洞的曲面,还有一个带一个洞的曲面(即环面),我求它们的和,我得到一个新曲面,它有三个洞.如果我对带三个洞的曲面和带一个洞的曲面求和,则得到有四个洞的曲面,诸如此类.我再说一遍,我所说的一切,都是指可定向的曲面.我们已经讲了曲面理论,这就是二维的情形.

现在我想转向三维的物体.

刚才,我们说到住在二维空间(就是曲面)上的人.他们非常小.他们看到的周边环境也很小,看起来像个平面.但他们可以提出问题:如果我们能看得很远,这个空间会是什么样子?那我们自己呢?我们也非常小,生活在三维的环境中.我们生活的环境是不是类似于三维球面的某种东西呢?如果我们能看得很远,会不会发现一个洞呢?你可以对二维情况提出这些问题,但三维问题更与我们切身相关.

我们看到一个三维空间,我们有了望远镜,而且它越来越高级.如果我们能看到足够远,我们会发现什么?我们住的空间是等价于球面呢,还是有一些洞.这是个严肃的问题,真可以对宇宙的性质提出这个问题.所以,如果你一定要得到对今天我所讲的一个物理解释,这就是了.

我从二维出发,是因为它比之三维情况更容易定义加法.

一男士:但裤子是三维的.

塞吉·朗:不,不!裤子的曲面是二维的.当然裤子一定是存在于三维空间中的,但裤子本身是二维的.你必须区别物体自身,即曲面本身的维数和它存在于其中的空间的维数.现在要考虑的物体,它本身就是三维的.

例如,取一个球,就是球面的内部,实心球.它是三维的.



球面



球

这个球,不考虑它的边界(即球面),它就等价于整个三维空间,如同圆盘的内部等价于平面 R^2,而没有边界的线段等价于直线 R.这里,字母 R 表示实直线,我在右上角加一个 2 表示维数为 2 的空间.对三维空间,我就写 R^3.

一维的东西是曲线,二维的东西是曲面,三维的东西该叫什么呢?

听众:体积(volume).

塞吉·朗:如果你喜欢的话,那也行.但"体积"一词有其他意思.它可以指这个空间的体积的数值.例如,手提箱的内部,你可以说它是体积,但你更可能说它的体积是 3 立方英尺.我们要对此加以区别.

在橡皮几何中,我们不去量体积的数值.因为一个东西可以等价于另一个大得多的某物,只要拉一拉,扯一扯就行.

我也可以说关于三维的物体,但它们在数学上有一个名字,相当专业的名字 —— 它们叫做流形,三维流形.我不喜欢这个名字,但大家都这么叫.同样,我也有橡皮等价的概念,我也有紧致流形的概念,换言之,在任何变形下都不会跑到无穷远去的流形.我还有边界的概念,边界应该是什么?

女士:曲面.

塞吉·朗:完全对.你已经懂得我所讲的了.

好了,我已经讲了一个小时.前两年的演讲,都是一个小时结束,然后提问题,听众又待了很久.因此,只要愿意都可离去,我们可以休息几分钟.我想讨论的主要课题是三维流形的分类,甚至涉及某些非紧致东西.

讲二维情形时,我给出了曲面分类定理:只有球面、环面、带越来越多的洞的曲面.在三维情形下,这是一个极端困难的问题,数学家们至今仍企图去解答.这正是瑟斯顿的贡献,他提出描述所有三维流形的一个猜测.这里也有求和,也有洞.但都要复杂得多.这就是我下面要讲的.

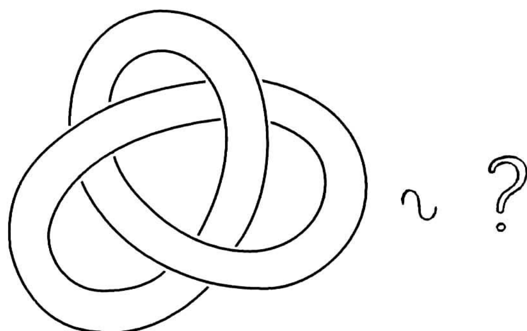
但现在,暂停或者休会,看你喜欢哪种修辞.

[掌声响起来,有人问,是否有足够时间到街对面喝点什么?我说,可以.约 15 分钟后继续讲]

2 第二小时演讲

[开始时,大厅座无虚席,约有230人.大约$\frac{3}{4}$的听众回到大厅听第二部分的演讲]

[黑板上,某位听众画了如下的图]



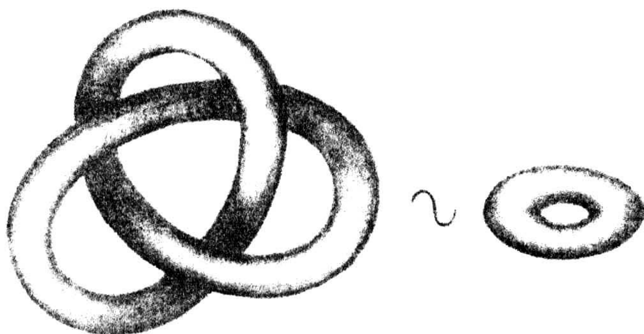
塞吉·朗:[看图]啊,画得很好.它像一个结,但有一个洞.对我前面所讲有没有问题?

男士:这个曲面是否等价于环面?

塞吉·朗:问得很好.你认为呢?

一听众:这个曲面有几个洞?

塞吉·朗:啊,这是有一个洞的曲面,放在三维空间中,而如果你想在三维空间中把它变形为环面,那是办不到的.但如果你允许在更高维空间中作变形,它可以变成环面.你可看到,有多种方式来表示同一曲面.



男士:如果有人在曲面的内部行走,那他无法判别是否有结.

塞吉·朗:是的,精彩的注解.它就像一个结.如果你在结上一直往前走,那你迟早会回到出发点,但你没法知道你不是在一个圆上走.



[有人举手]

塞吉·朗:请.

一女士:麦比乌斯(Möbius)带是怎么回事?

塞吉·朗:我已经说过,我只想讨论可定向的曲面,就是要避免这种东西,因为我想避开过多的技术性而谈简单一些的命题.提到可定向性只是为了防止有人抱怨我说得不全面.如果我涉及不可定向曲面,我就没时间谈三维问题了,但这恰恰是我想谈的.好了,麦比乌斯带,也许你们中不少人已听说过,而现在我没时间去谈了.你们大概对三维的情形知之甚少.

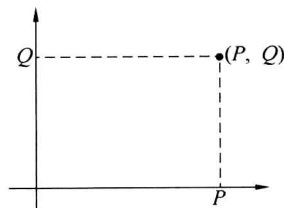
此外,三维情形与我们生活的世界密切相关.我早已说过,数学家研究各种可能性,各种模式.作为数学家,我们对这些模式之美妙感兴趣,而不在乎其物理应用.今天,我已经对曲面进行了分类,我还对三维流形的分类感兴趣.我试图对它们进行描述.一旦我们知道了所有的三维流形,我们自然会问,其中哪一个对应着物理世界 —— 我们生活于其中的世界.物理学家要研究这些模式,以便找出哪些能与现实世界相吻合.我本人不做物理,数学与经验世界的关系常让我搞不清.打从我当学生以来,我总感觉我对物理无能为力,物理真的引不起我的兴趣.

一女士:没错,学生们要把所学的数学知识应用于物理是很困难的.

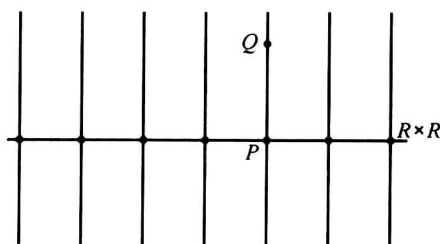
塞吉·朗:我不必隐藏我自己的品味,但我也不把它看成自然规律.我喜欢对事物进行分类,我构造一些模式,然后对物理学家说:“请选择适用于你的一个模式吧!”另一方面,也有其他数学家,像阿梯亚或者辛格等,他们对物理感兴趣.反过来,也有物理学家对数学有深刻理解,既搞数学也搞物理.我是不行了.但我向学生们讲得很清楚,如果他们喜欢,也有此能力,那我鼓励他们二者都搞.但任何人都有其自身的限制.

好,让我们回到三维情形.要画图变得困难多了,因为举例来说,就是三维球面,我也不能画得很清楚.通常的球面,二维球面,我可以把它看做与一个给定点,称为球心,保持恒定距离的点的集合.而三维球面 S^3,也可以定义为在四维空间中与球心保持恒定距离的点的集合.我们虽然画不出来,但我们相信有这个东西.

我所能做的,是画一些图,它可以帮助你设想会发生什么.或者,做一些别的表示法.举例来说,在平面上,画两条坐标轴,和两轴上的点 P 和 Q.



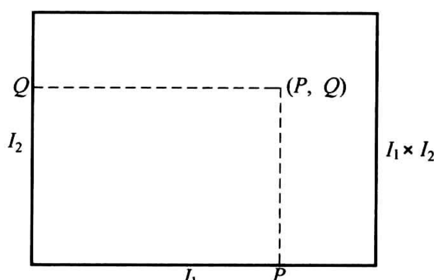
我可以写点(P,Q)，其中Q在直线R上，而(P,Q)在平面上．这种做法称为乘积．这就像在每一个点P上放一条直线，而点Q就在这条直线上游走．



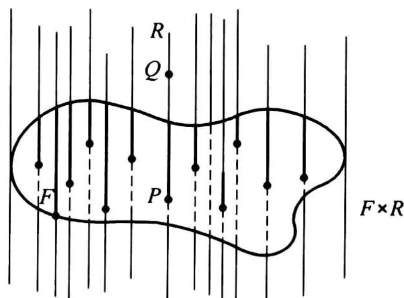
我已经说了，这种做法称为作乘积．我们看到，平面是直线R和直线R的乘积，我们把它写成

$$R^2 = R \times R$$

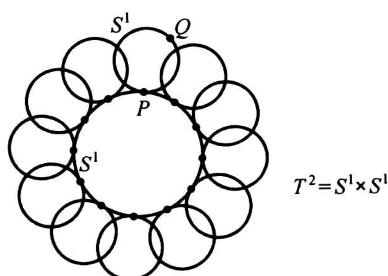
类似地，如果我有两条线段I_1和I_2，而我取第一条线段上的所有点P的集合和第二条上所有点Q的集合，那所有这些点对(P,Q)组成了长方形上的所有点，就像这样



对任意两个集合，我可以用此法作出乘积来．如果我有一个二维的曲面F，我可以把它和任何一维的东西作乘积．如果F是二维的，那这个乘积是三维流形．它是一些点(P,Q)的集合，其中P是曲面上的一个点，而Q属于一个一维空间．



让我再举一个例子．设S^1是圆．我作乘积$S^1 \times S^1$，即所有偶(P,Q)的集合，其中P在一个圆上，而Q在另一个圆上．对第一个圆上的每个点P，我可以让第二个圆上的每个点与之相配．



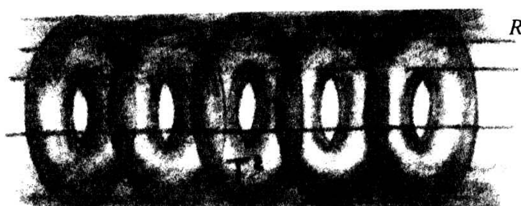
我会得到什么曲面?

听众:环面.

塞吉·朗:很好,一个环面.$S^1\times S^1$ 是一个环面.我把它称为 T^2,T 是环面(torus)的第一个字母,而 2 表示它是二维的.

乘积的概念使我可以构造高维的东西,我可以把它写出来,而不必画出来.

现在,我可以得到更高维的东西.举个例子,我作 T^2 与 R 的乘积,环面与直线的乘积.画出它来很困难,但我可以这样来表示它,画一个环面,一条直线,像这样,我可以把环面想象成这个东西,三维的乘积的一个截面.



当然,这个图并非一个逼真的图,但它可以帮助你想象.

现在我要画更复杂的三维物体.我已经有了三维球面,有了乘积 $T^2\times R$,但我想要某种相当于有洞曲面的东西.这种东西会是什么样子?

一女士:像一根壁很厚的管子.

塞吉·朗:对的.我想要有洞的,环面,还有跑到无穷远的东西.[塞吉·朗画出下图]



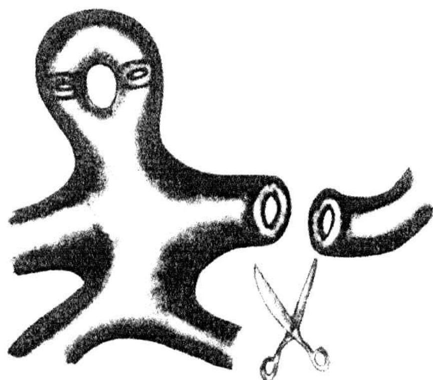
[笑声]好,这是个三维的东西,它该叫什么呢?

听众:章鱼.

塞吉·朗:很准确,章鱼.二维情况,我有裤子;三维情况,我有章鱼.这样一来……设若我有一把剪刀,我剪掉一条腿,我得到什么?

听众:……

塞吉·朗:我们这个章鱼没有边界①.如果我切去它的腿,我得到某种东西,它的边界是二维的,会是一个环面.

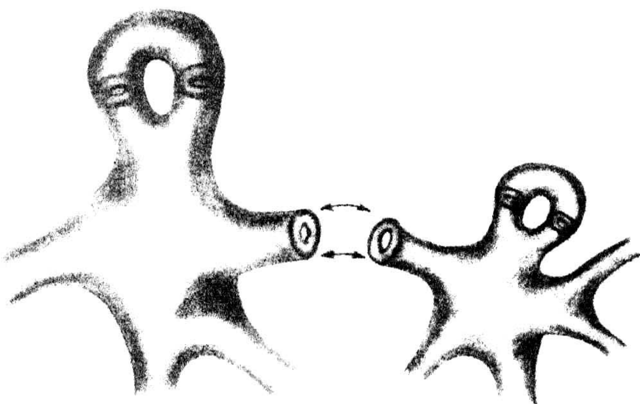


现在,设你有某种三维东西,它的边界是一个环面,你还有另一件东西,它的边界也是环面.你一定情不自禁地想对它们做点什么.你想做什么?你[指向一女士].

女士:把它们粘到一起.

塞吉·朗:正确,就像早先对圆所作的.所以我取两个章鱼.

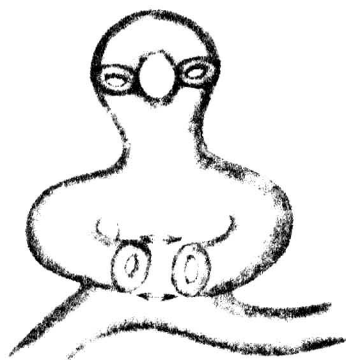
把每个章鱼切下一条腿,我得到两个环面,我把它们粘到一起.这样,我们对两个章鱼沿着环面求和了.



① 不幸的是,不可能逼真地画下此图,还有三维球面 S^3 也如此.按我们的作图法,画了一个边界,但它总算大体能表现事情的真相.



我也可以对一个章鱼作这种手术,切下它的两条腿,把所得的两个环面粘起来.



先前我们还有盖子.现在我们有什么呢?我有一个环面为边界,我想消去边界,怎么办?

一男士:粘一个戒指上去.

塞吉·朗:很好,那是对的.戒指是环面的内部.我取一个戒指,把它粘到环面上,这与在圆上粘一个盖子类似,但是维数不一样.这样,我就可以消去一片边界.

男士:一条章鱼可以有多少条腿?

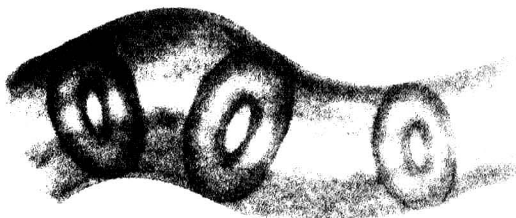
塞吉·朗:随便多少都行.两条章鱼的腿的数目还可以不一样,这种情况下它们就不等价.[笑声]

男士:如果一条章鱼的腿的数目是奇数,另一条章鱼的腿的数目是偶数,你怎样作它们的和呢?

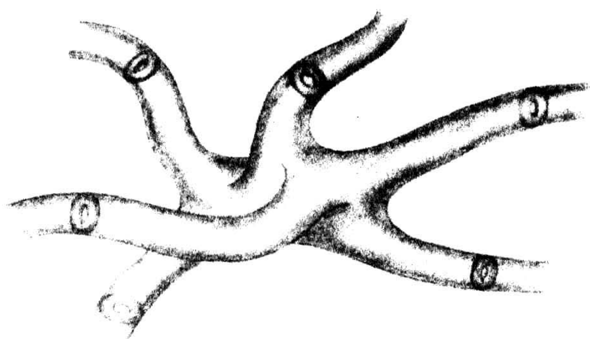
塞吉·朗:我没有说要你把一条章鱼的所有腿与另一条的所有腿都粘上.你可以把一部分腿粘起来,把另一些腿用戒指盖上.

男士:什么是 $T^2 \times R$?

塞吉·朗:好,$T^2 \times R$,它是……嗯……它像一个没有洞的章鱼,只有两条腿.



如果我切开 $T^2 \times R$,作成一个截面,我得到一个边界,它是一个环面.也存在没有洞的章鱼,但有很多条腿.



就像对曲面一样，我们说一个章鱼是不可约的，如果你把它表成两个章鱼的和，那其中之一必定等价于 $T^2 \times R$，或者就是盖上一个盖子，就像那位男士刚才所说，粘上一个戒指.

如果我取一个章鱼，求它与 $T^2 \times R$ 的和，那我得到一个新的章鱼，它与原来的章鱼是等价的.从等价的观点来看，什么也没改变.我们说，对于这种切去和粘上腿的办法作成的加法来说，$T^2 \times R$ 是零元.

现在完全可以相信，经过有限步这种加法我可以去掉所有的腿.让我把它写下来.

> 经有限步加法后，可以有多种方法消去所有的腿.这样，我们得到一个三维流形，它是紧致的，没有边界.

女士：但可能有洞.

塞吉·朗：是的，肯定是的.我们消去了腿，但创造了洞，而且可能有很多洞.这是构作三维流形的方法之一，这些流形是紧致的，没有边界……还有，是可定向的，加上这条足以使我良心上过得去，也使别人无法抱怨.

当然，下一次你去海滩，你可以试试——[笑声]切下章鱼的一些腿，然后粘到一起，甚至可以先打一个结再粘.

为了对章鱼分类，我们必须对不可约的章鱼进行分类，然后对加的方法进行分类，就是切腿和粘上的方法.直到现在，我已经描述了一些几何模式；先是曲面的模式，然后是三维流形的模式，章鱼，和三维球面 S^3，它不是章鱼而是另外的东西.

一听众：它没有洞.

塞吉·朗：是的，它没有洞.于是人们可以提出这样的问题.

取所有的紧致三维流形，没有洞，也没有边界.你能把它们描述清楚吗？这个问题至今未能解决.庞加莱猜想说：紧致的、没有洞的、没有边界的三维流形





一定是三维球面 S^3. 他猜测没有其他东西. 当然,应该把"没有洞"的意思说得更精确一些,但今天让我们把这个技术性问题放在一边.

很多人企图找到庞加莱猜想的答案,但至今无人成功. 1960年,斯麦尔对五维及五维以上的情形证明了类似的猜测. 从那以后,就只剩下三维和四维的情形了. 但维数愈小,证明愈困难,因为你没有足够的空间左右移动. 1981年,弗里德曼(Freedman)解决了四维的情形. 很多数学家的工作对此都有贡献. 他们用"纯粹的思想",而不是用很多复杂技术,把这门理论尽力发展,终于攻破难关.

弗里德曼经过6至10年的努力得到成功. 这是非常困难的,很带技巧性的、极其复杂的工作. 它是现代数学最伟大的成就之一,是一流水平的工作.

现在只剩下三维的情况了.

因此,我说不出三维流形的完全分类,因为庞加莱猜想还没证明. 除庞加莱猜想,对其他的三维流形,瑟斯顿提出了一个猜测(而他自己已经证明了相当大一部分):

可以列出一张具体的某些流形的表,不算太大,使得:

> 每一个没有边界,可定向的,紧致的三维流形或者是在这张表中,或者是一些章鱼的和.

至此,我已经把涉及橡皮几何的部分讲完了,至于瑟斯顿猜测和它的那张表的更确切的描述,则需要完全不同的另外一些思想.

相当有趣的,甚至是非常有趣的,是那张表中的流形的构作方法与构作带腿的章鱼的方法差不多. 换言之,用同一种办法可以造出有腿的和没腿的流形. 要做这件事,我们要放下橡皮几何,做完全不同的另一种几何. 多数听众可能已经听说过非欧几何,但我们是要做三维的非欧几何.

继续讲下去之前,谁有问题?你们对今天所讲感觉如何?

一男士:在章鱼身上的每一点,有一个环面还是有多个环面?

塞吉・朗:那不一定. 如果我切下一条腿得到一个断面,那我得出一个环面. 如果我切另外的什么地方,那就不一定了. 只有切的地方对了,切的是腿,才能得到一个环面,数学家会这样说:

> 章鱼是一个非紧致、无边界的三维流形,它有有限多个端头,每个端头等价于 $T^2 \times R$.

所以,如果我在某点切章鱼,切下的不是腿,我切的某处不是等价于 $T^2 \times R$ 的端头,那就无法保证得出一个环面. 其实,如果我切某处,我可以切一个球,就

像你取一个冰淇淋球,你会留下一个球面为边界.你也可以想象瑞士奶酪中的气泡,与曲面的情形一样.在曲面的情形,我切下一个圆盘,留下一个圆作为边界,那正是我们定义用粘接方法求两个曲面之和时所为.

切下一些球,我可以把一些三维流形加在一起.在第一个流形上切下一个球,再在另一流形上切下一个球,每个流形上都留下了边界,它是一个球面.把这两个球面粘到一起就得到两个流形的和.说一个流形不可约是指,若你把它表为两个流形之和,则其中之一必等价于一个球面 $S^3$①.米尔诺(Milnor)1962年证明每个紧致、无边界三维流形可以表示成不可约流形之和,且表示法本质上是唯一的②.这把三维流形的分类问题化成了不可约流形的分类问题.当然,总是在等价的意义下谈的.

有其他问题吗?没有?好,那我们继续下去,我们来谈距离几何,以及非欧几何.但我已讲了两个半小时.我该拿这些非欧玩意儿怎么办呢?你们想走了吗?听够了吗?你们说怎么办就怎么办吧.

一女士:不,我们留下,你诱发了我们的好奇心,我们愿听下去.

塞吉·朗:啊,我诱发了你的好奇心!那么,章鱼沉下去了.好.[笑],休息五分钟如何?

听众:不,我们都坐在这里,请继续吧.

一男士:现在我们正想听,继续吧.[这位男士做了一个手势,意指"继续"]

塞吉·朗:好,那就继续讲.你们对数学还真有胃口!如果哪位想走,有约会,请便.[笑声]不是我想赶你走,但是……

[若干人离场,在最后一小时演讲中,陆续有人离场]

3 第三小时演讲

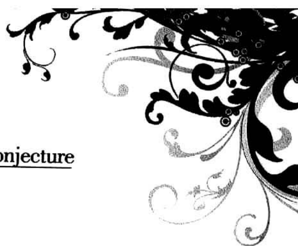
现在,我们放下橡皮几何,来谈距离几何.在实直线上,或者平面上,或者通常的三维空间中,我们都有距离的概念.我们现在感兴趣于一种新的等价关系,它得保持距离.我们把保持距离的变换称为运动.在欧氏几何中有运动,在非欧几何中也有运动.但我想从欧氏情况开始,以求讲清其中的思想.运用这些运动,我们可以规定某种同一,它会使我们再次看到章鱼,而距离几何就这样与橡皮几何汇合到一起.所以,我们现在要讲的东西,是相当实质性的.

让我们从直线 R 开始,它上面有 $1,2,3,\cdots,-1,-2,-3,\cdots$

① 注意这里的"不可约"是对沿球面作和而言,与前面的沿环面作和而言的不可约是不同的.这是两种不同的求和法,提到求和时应在上下文中指明是指何种求和法.

② J. Milnor, "A unique factorization theorem for 3-manifolds." Amer. J. Match. 84(1962).







如果给定了一个方向,和一段距离,我把它表示成一个箭头:



取一个点 P. 我可以让它沿箭头所指方向移动箭头给定的距离. 于是我得到一个点 Q,把它称为 P 的平移并记成 $\tau(P)$①

$$P \to Q = \tau(P)$$

为具体起见,取箭头的长度为 1. 于是 1 的平移是 2;2 的平移是 3,等等. 现在我把一点 P 和它的所有平移同一起来. 让我来画一个点和它的平移.



如果我作这样的同一,我得到一个圆,就如同本讲开头所作的. 我取一条线段,带有端点,而我让两个端点等同起来,就得到一个圆.



对你来说,期望作这样的同一,是很正常的反应,很数学化的反应. 我们现在就正在用这种方法. 此法在定义橡皮几何中的等价性时是不允许的. 简而言之,在一条直线上,我让每个点与其按给定方向、给定距离的平移同一起来,就得到一个圆.

当然,如果 P 是一个点,那我把 P 和下一个点 $Q = \tau(P)$ 同一,然后,再下一个 $R = \tau(Q)$. 我应当怎样写 R?

听众:$\tau(\tau(P))$.

塞吉 · 朗:对了,$\tau(\tau(P))$,也可写成 $\tau^2(P)$. 如果我再来一次,我就写

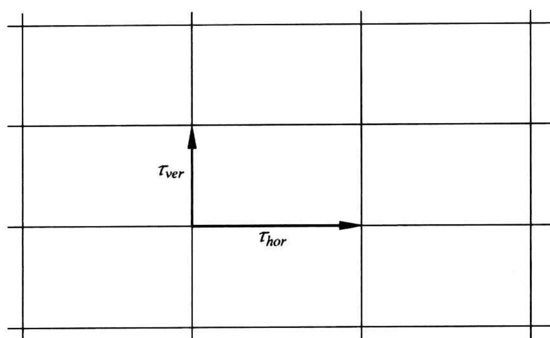
$$\tau(R) = \tau(\tau(\tau(P))) = \tau^3(P)$$

如果往相反方向作平移,我就写 $\tau^{-1}(P)$.

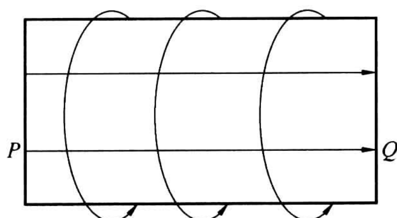


好了,我们转向二维情形. 这时我有水平的平移和垂直的平移,分别表示为 τ_{hor} 和 τ_{ver}.

① 字母 τ 是希腊字母,发音类似于 /tau/. 我本来也可以用字母 T,但我们已经用 T 表示环面,所以换一个字母.



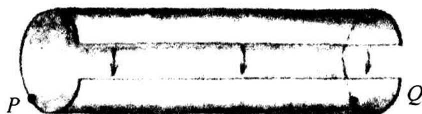
我现在又可以往两个方向作平移和同一:水平的和垂直的方向.如果我把下图中点 P 和点 Q 同一起来,就是把长方形的左边和右边同一起来.我还把上边和下边同一起来.



我得到什么?

女士:一个球面.

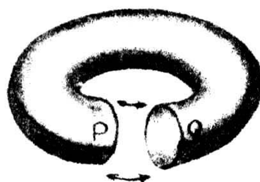
塞吉·朗:不!请看,同一的意思是什么?当我把上边和下边同一起来,我得到柱面.



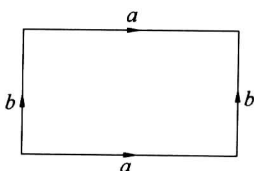
现在若把两边同一起来,我得到什么?

听众:环面.

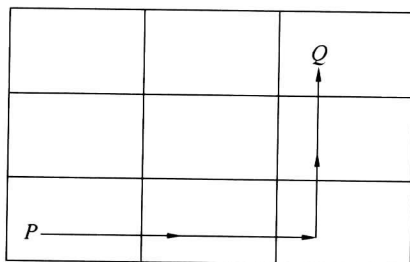
塞吉·朗:对了,一个环面,T^2.



现在你看到了,我可以由二维的图形,经过水平的和垂直的同一而描述环面.



我可以对整个平面做这种同一.



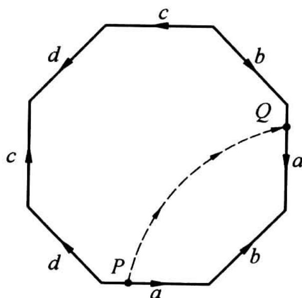
如果平面上的一个点可以经过若干次垂直的或水平的平移跑到另一个点，我就说这两个点等价.但这种等价与橡皮几何中的等价是不同的.这里我加进了方向和距离的概念.现在我需要这两个概念，而从前我完全没有管它们.

所以我必须说清楚，所指的是哪一种等价.我需要两个不同的词来表示这两种等价.我得约定某种术语，我马上要来系统地解说它.

好，运用作同一的办法我得到一个环面，就是有一个洞的曲面.如果我想要得到一个有多个洞的曲面，你能猜出该怎样做吗?这里，我从一个长方形出发，作出了一个环面.如果我想要得到一个，例如说，有两个洞的曲面，我应该怎样作同一呢?

女士:画另一条直线在当中，诸如此类.

塞吉·朗:是的，你说对了，应该画更多的线，但你说得也不完全对.让我告诉你怎么做.不用四边形，而用一个八边形.



像我画的那样去做同一.例如，把我画的点 P 和点 Q 同一起来.

如果我想要一个有三个洞的曲面呢?

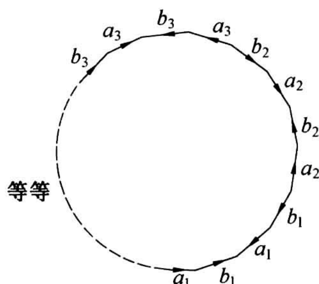
听众:用十二边形.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

塞吉・朗:对,两个洞的曲面要八条边,三个洞的曲面要十二条边.如果我想要 n 个洞的曲面呢,我需要 ——

听众:$4n$ 条边.

塞吉・朗:很好,我把它画成这样.



这就是如何得到一个有多个洞的曲面的方法.

男士:如果你有一个六边形呢?

塞吉・朗:那我不能像前面一样得出一个有洞的曲面.我可以得到某种别的东西,不可定向的曲面,但是,今天,我只想谈可定向的曲面①.但你的问题表明你已经懂了我所讲的.

你看到了,环面可以作为平面的商而得,就是说用取同一的方法,这将被写成一个斜线

$$T^2 \sim \text{同一} \setminus R^2$$

这里的同一是平移.

男士:2 的意思是什么?

塞吉・朗:它表示维数.写在肩上的数字,即上标,就是用来表示维数的,不表示其他意思.我发了誓不用数字,但是肩上写一个小小的 2 仍然是有用的.我说了我只谈几何的事情,但是这个 2 只表示维数.你们允许我这样做吗?

听众:可以.

塞吉・朗:谢谢.因为我答应了今天不用数字,但这个 2 不是真正的数字.[笑声]

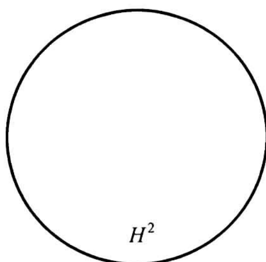
好,所以我把 T^2 表成了 R^2 关于平移的商,而这是欧氏空间及其上的平移.现在,让我转向非欧几何的运动.

非欧平面的模式之一是圆盘.我把它称为 H^2,H 代表双曲,H 是双曲的英文词(hyperbolic)的第一个字母.

---

① 这依赖于作同一的边的位置和方向.有时会得到环面,有时会得到不可定向曲面.这是一个很好的练习:研究把 $2n$ 个边的多边形作同一能得到哪些曲面.

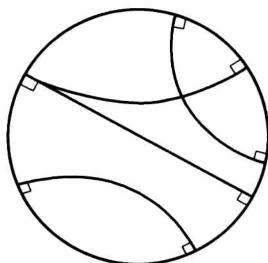




我需要双曲距离的概念，和在这种距离下的"直线"的概念．今天的听众中，一定有些人已经知道这些．谁知道？

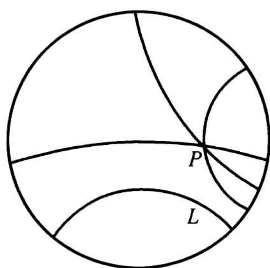
听众：？？？

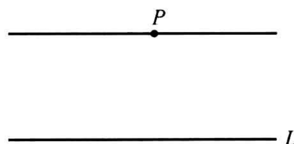
塞吉·朗：好，我来讲它的意思．按定义，双曲直线就是 H^2 中的一段圆弧，与边界相垂直者．我可以把它画成这样，这里是一些双曲直线．



其中有些互相相交，而另一些则不然．垂直一词的含义与欧氏几何的一样．

你看到了，你可以有无穷多条直线通过一个点 P，它们都不与另一条直线 L 相交．这在欧氏几何中是不会发生的．



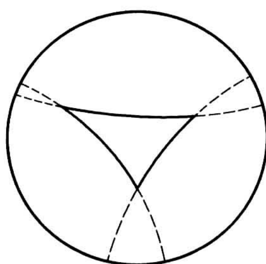


非欧情形　　　　欧氏情形

在欧氏几何中，给定一条直线 L 和一个点 P，一定正好有一条通过 P 的直线与直线 L 相平行．

我们定义三角形，就像在欧氏几何中那样．这是一个三角形，它的边是线段．

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$



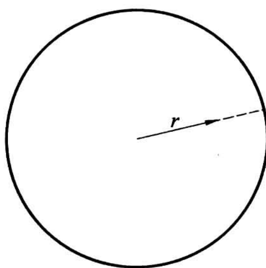
现在,我们已经看到直线是什么样子,我们来描述双曲距离的概念,以及从这种观点看来,空间是什么样子.然后,我们会看到它和章鱼的关系,与三维流形分类的关系.我想讲到瑟斯顿猜测为止.

所以,我们得定义一种新的距离,称为双曲距离.它又被法国人称为庞加莱距离,而被俄罗斯人称为罗巴切夫斯基距离.我把它称为双曲距离,两面都不得罪.[笑声]

要完全描述双曲距离,我需要一些公式,而我不想在这里写一些太技术性的公式.但我要说到当我从中心往边上走时距离变化的速度.这意味着,如果 r 表示从中心算起的欧氏距离,那么,双曲距离沿着一条射线变化的速度可以很简单地写出来,那就是

$$\frac{1}{1-r^2}$$

这里,我假定了圆的半径是 1.所以,如果我从圆的中心出发,沿一条射线走向边界,变化的速度如何呢?你看,如果 r 接近于1,那 r^2 也接近1,而 $1-r^2$ 就很小,于是,这个分式就很大.因此,当我走近边界时,距离变化的速度就很大,越来越大.有一个用来表示距离的公式,它要用到对数.有谁听说过对数吗?

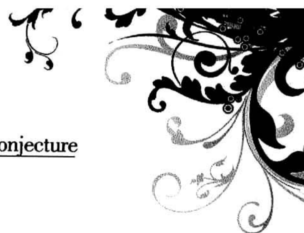


[很多人举手]

啊!很多人知道.那么,我来写出这个公式,不懂对数的听众可以不听.沿着射线的双曲距离是

$$d=\frac{1}{2}\log\frac{1+r}{1-r}$$



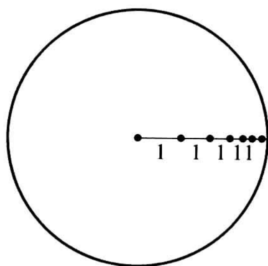


所以距离随着我们接近边界而越来越大.

你看,这很像欧氏几何的情况,也很像我们的世界.如果我们从中心出发,向边界走去,尽量走快些,那么,在我们这个宇宙,会发生什么情况呢?

我们知道欧氏模式不灵了,我们知道空间会变得弯曲,有一些像双曲直线.我们来加快速度.假设一束光线往同一方向射去.我测量它的速度,那是每秒300 000千米.现在,假定我走得快一些,如果世界是欧氏模式,那我再来测量,我会发现光速的值减小了.是吗?好了,答案是不,完全不.我总是测得同一数值.如果两列火车用同一速度往同一方向开,那么它们彼此之间相对就没有动.但是,这不适用于光,光速是恒定的.原因就是,如果我走得越来越快,我就变得越来越小,而我用以测量的工具,也变得越来越小,我用它来测量光速,总得到同一数值.

在双曲平面中,我们碰到同样的现象.这里我画了一些点,彼此距离是1个单位,它们排列在一条射线上.



所以,我们怎么能知道我们不是住在这样的空间里呢?我们走得越远,我们对另一边的事物知之越少 —— 甚至"另一边"是否有意义都不知道.

但是,我们依然会提出问题:我们住的世界是什么样的?那么,数学家创造一些模式,而物理学家来判定哪一种模式适合于我们住的世界."另一边"的含义是不清楚的.若用另一种方法来看双曲平面,不把它放在通常平面中,而是内在地看它,它就是它自己,那就无所谓"另一边".一个可以提出的问题是,我们的宇宙是不是放在另一个宇宙之中的.但我们没法同这个大宇宙打交道,我们只能从它对我们的宇宙的影响来推测它的性质.

好了,让我们回到数学.我有了这种模式,我可以做同一,就像在欧氏模式中一样.

[有人举手]

塞吉·朗:请.

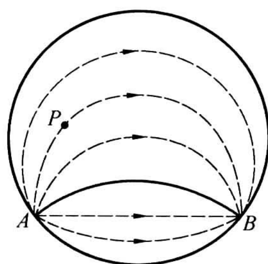
男士:你定义了到中心的距离,但能不能也定义任意两点间的距离呢?

塞吉·朗:当然可以,但要这样做会涉及更多技术性问题,公式也更加复杂,需要用双曲函数来写这些公式,所以,我不想去做了.

好，现在，我想作同一.我需要某种运动，这种运动应该保持双曲距离.像从前一样，我可以定义平移.如果我有一条双曲直线，它给了我一个方向，我还有双曲距离的概念.取任何一个点 P，现在问：哪一点是它沿这个方向的平移？点 P 沿着哪条曲线来运动？

听众：???

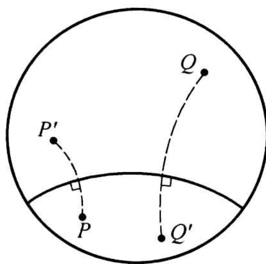
塞吉·朗：设 A，B 是双曲直线的两个端点，就像下图中所画.点 P 在这条直线方向上的平移会沿弧 APB 而运动.



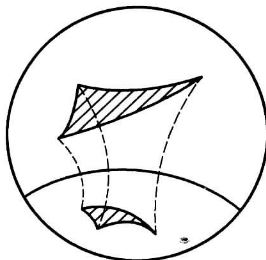
我可以沿一个方向平移，也可以沿反方向平移，也还可以多次平移，等等.平移是保持距离的运动的一个例子.还有另外的例子，你知道是什么吗？

一高中学生：旋转，取镜象.

塞吉·朗：完全正确.在双曲平面上，旋转与欧氏平面上一样.至于取镜象，我来画一个点 P 及其关于一条直线的镜象.



我还可以画三角形，以及三角形关于这条直线的镜象.







一个点关于中心的镜象与欧氏情况是完全一样的.

女士:那双曲平面有一个中心,是吗?

塞吉·朗:不,双曲平面本身没有中心,但我们作的这个模式有中心.在双曲平面上,到处看起来都是一样的,不管这地方按欧氏距离与边界有多近.给定两个点 P 和 Q 总存在一个平移把 P 送到 Q.可以说,双曲平面是各向同性的.在双曲平面上,你离边界总是无限远.当我用圆盘来表示它时,我取了一个中心点,也就像给欧氏平面选了一个坐标原点.

好了,让我们再来作同一.我有旋转、平移,以及取镜象,我还可以把它们连起来作,重复地做.一般来说,这些就给出了所有的保持距离的变换.如我先前所说,我把它们简单地称为运动.

现在,我还要定义另一概念,运动群.我将说:

> Γ 是一个运动群是指:
>
> 1) 当两个运动 M_1 和 M_2 在 Γ 中,那它们的合成 M_1M_2 也在 Γ 中.
>
> 2) Γ 中一个运动 M 的逆运动也在 Γ 中.

这里说的合成 M_1M_2 是这样一个运动,当你把它作用到一点 P 时,你得到 $M_1(M_2(P))$.一个运动的逆运动是指:若运动把 P 送到 Q,则逆运动把 Q 送到 P.所以现在我有了运动群的概念.

我还需要离散群的概念.让我们从一个例子说起,在通常平面上,平移作为要考察的运动.取一个点 P,然后平移它.



学生:这些点完全不相同.

塞吉·朗:是的.如果我取平面上一个有界区域,会发生什么事?

学生:这些点迟早会跑出这个区域.

塞吉·朗:对了.我可以对一个群做这件事.我说两个点 P 和 Q 关于群 Γ 等价是指存在 Γ 中的一个运动 M 使得

$$Q = M(P)$$

这是一个等价关系.我说 Γ 是离散的是指:给定一个点 P,那对 Γ 中所有可能的运动 M,点 $M(P)$ 中只有有限多个可以在空间中的一个有界部分中.这其实就是说在空间中任何有限部分中,在任何有界区域中,只能有有限多个点与点 P 关于 Γ 等价.

为了区分这种新等价与橡皮等价,我只得把 Γ 明确说出,所以我会简单地说 Γ - 等价.

现在,设 Γ 是离散的.我可以来作关于 Γ 的同一.我可以把所有相互 Γ - 等价的点同一起来.经过这种同一,我得到一个新空间.我把这个新空间写成

$$\Gamma \setminus H^2$$

这个空间也是二维的.

我刚才做了二维空间中的同一.当然,我也可以在三维空间中做.在三维空间中,我们取什么东西来作双曲空间的模式?

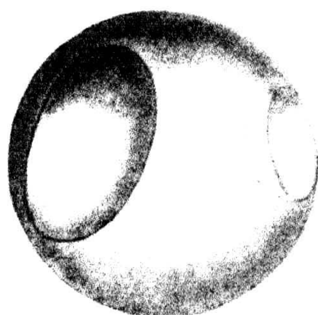
男士:球面.

塞吉·朗:是的,是球面的内部,一个球.在三维情形,我们有 H^3,它是通常的球,但是用双曲距离,这与平面上相似.当你跑向边界时,距离会变得任意大.

在三维双曲空间中,平面是什么样子?

男士:球面的一部分.

塞吉·朗:对了.



我们也可以定义平移,取镜象等.

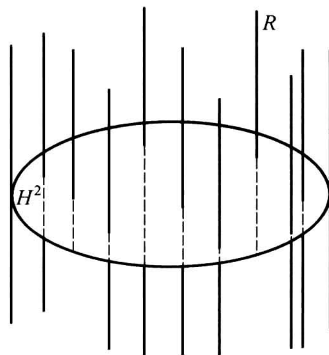
但是还有另一种办法来构作三维空间,就是用某个一维的东西和某个二维的东西来做.前面我已经用到了作乘积.现在谁能给我三维空间的另一个例子,它可以用做乘积的方法做出?

学生:取一条直线和一个双曲平面.

塞吉·朗:啊!非常好!那正是我期望你们想到的.所以,我们有了另一个例子,取双曲平面 H^2 和直线 R 的乘积,它可以写成

$$H^2 \times R$$

这个空间在 H^2 上有双曲距离,而在 R 上有通常距离.

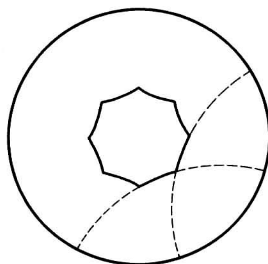




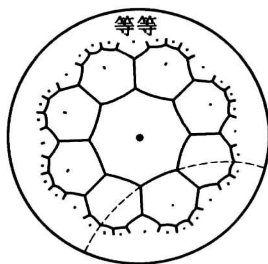
现在我们有了两个基本例子

$$H^3 \text{ 和 } H^2 \times R$$

我们已经做好准备,下面要与橡皮几何和橡皮等价联系起来.我首先要回顾一下曲面理论的一个经典定理.让我们回到多边形,不过这次是双曲平面上的多边形.它的边是双曲直线的一段,但这样的多边形与通常的多边形是等价的,就是通过多边形的边的同一来造出曲面的那种多边形.



双曲平面可以用这种多边形的平移来覆盖,也就是说,两个平移要么没有共同点,要么有一个边全是公共点.



这就像在欧氏平面上你可以用正方形,长方形来覆盖一样.但欧氏平面不能用正八边形来覆盖,而你却可以用任何正多边形来覆盖双曲平面.

如我们先前所做的,让某些边同一,现在变成按一个平移群来做同一.我们有定理:

设 F 是紧致、可定向、无边界的曲面,不等价于球面或者环面,则存在一个离散群 Γ 使得 F 等价于双曲平面 H^2 关于 Γ 作同一.换言之

$$F \sim \Gamma \setminus H^2$$

这个定理是 19 世纪发现的.但在瑟斯顿之前,没有人想到在三维情形下有与之类似的东西.瑟斯顿的重大发现就是猜测到应该有类似的结果,并且对某些情形证明了这个猜测.首先,我想谈到一个结果,它把演讲的第一部分——橡皮几何,与第二部分——非欧几何联系起来.

我们用 Γ 表示一个离散运动群,但我们进一步假定:

——对每个点 P,Γ 中的使 $M(P)=P$ 的运动 M 只能是恒等运动,即使每个点都不动的运动.

——Γ 中的运动保持方向.

我们永远假定 Γ 满足这两个附加条件,即使我没有明显提到①.

例如,设 Γ 是 H^3 上的一个运动群.那我们可以有两种情况.

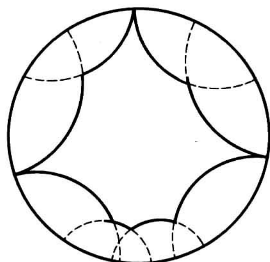
第一种情况:$\Gamma \setminus H^3$ 是紧致的.

在这种情况下,关于 Γ 作同一而得的空间是紧致的、无边界的三维流形.这是得到这种流形的方法之一.

第二种情况:$\Gamma \setminus H^3$ 不紧致.

在这种情况下,我们不单要用到距离的概念,而且要用到由它导出的体积的概念.当我作关于 Γ 的同一之后,有可能空间 $\Gamma \setminus H^3$ 的体积有限.以下总假定 Γ 表示一个群使得 $\Gamma \setminus H^3$ 的体积有限.

当然,在二维情形也有类似的现象.可能有的多边形之边跑到边界上,因此这多边形的顶点离中心任意远.



但你可以有一个运动群 Γ,甚至是一个平移群,使得你关于 Γ 作同一而得的曲面有可以跑到无穷远的顶点.

同样的事也会在三维情形中发生,但这时就很难画出来了.跑向无穷远的那部分像根管子.

除 H^3 外,也可以取 $H^2 \times R$,然后考虑群 Γ 使得 $\Gamma \setminus (H^2 \times R)$ 是紧致的,或者是非紧致的但体积有限,它的一些管子可跑到无穷远.这些管子看起来像什么样子?

定理 设 Γ 是 H^3 或者 $H^3 \times R$ 上的离散运动群,让我们作关于 Γ

① 平移其实就是这种情况,而这些条件将排除 Γ 中包含取镜象的可能性.





的同一.设所得空间的体积有限,那么.这个空间或者是紧致的,或者是一个章鱼.而且,$\Gamma \setminus H^3$ 是不可约的章鱼.

在这个定理中,应把群 Γ 看成满足上述附加条件,例如,$\Gamma \setminus H^3$ 或 $\Gamma \setminus (H^2 \times R)$ 的体积是有限的.

一女士:是离散的,再附加这些条件吗?

塞吉·朗:是的,这是一个必须的附加条件.我要假定,经过作同一,所得的流形体积有限.

还有一个逆定理,它已经给出关于章鱼分类的某些思想.

定理 每一个不可约章鱼等价于

$$\Gamma \setminus H^3 \text{ 或 } \Gamma \setminus (H^2 \times R)$$

型空间.

所以,我们又回到了章鱼!很意外.我们从一个完全不同的观点出发,我们在一个有距离的几何中作同一,取保持距离的运动,我们得到什么?紧致流形和章鱼!这就是关于橡皮几何的第一部分和关于距离几何的第二部分的关联.

现在来看瑟斯顿的猜测.我们看到两个例子,H^3 及 $H^2 \times R$,都是有距离的空间.我曾经提到适当定义的空间的表.它由

$$R^3, S^3, S^2 \times R, H^3, H^2 \times R$$

这5个空间,以及另外3个空间组成.而那3个我不想去写,因为它们太带技术性了.一共是8个空间.让我们用 X 表示这8个空间之一.

现在瑟斯顿猜测可以作如此陈述:

猜测 设 V 是三维流形,紧致,无边界,以及可定向.(这使我们避免事情太复杂)设 V 对于沿球面求和是不可约的.那么,V 等价于下列情况之一.

—— 存在唯一的空间 X,它是8者之一,及一个群 Γ,使得 $V \sim \Gamma \setminus X$ 且是紧致的.

——V 是一些章鱼的有限和,且每个章鱼等价于某个 $\Gamma \setminus X$,其中 $X = H^3$ 或者 $X = H^2 \times R$.

除此之外,在第二种情况中,还有某种唯一性.或多或少地,这意味着如果

我们把 V 写成章鱼的最小和,那么,出现在这个和中的 $\Gamma \setminus x$ 本质上是唯一决定的,即在某种等价意义下是唯一的.把这点说精确要用到太多术语,要把"本质上"的意义说清楚也是这样.你必须定义一些新的等价,现在不是做这些东西的时候.

所以你看到,要得到章鱼,你所需要的就是 H^3 及 $H^2 \times R$.这就是瑟斯顿企图去证明的定理,而他其实已经证明了相当一部分.①

男士:庞加莱猜想与此有关吗?

塞吉·朗:S^3 就在这张表中,要得到 S^3,Γ 就是平凡群.庞加莱猜想在这张表中就这样孤立出来,你拿它没办法.

女士:我有点跟不上了,R^3 与 S^3 的区别是什么?

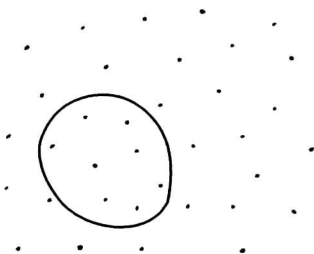
塞吉·朗:R^3 不是紧致的,它是我们周围的通常的欧氏空间,而 S^3 像一个球面,它是紧致的,但 R^3 跑到无穷.

[有人举手]

塞吉·朗:请.

女士:你可以重复一下离散的定义吗?

塞吉·朗:"离散"意味着,如果 P 是任一点,我们用 Γ 中所有可能的运动 M 去作用在 P 上,也就是对 Γ 中的所有 M 来考察点 $M(P)$,那么,在空间的任何有界区域中只有有限多个这种点.



这就是它的意思,Γ 是离散群的意思.再加上我所作的附加条件.

一高中学生:如果你取一个群,它不是离散的,你会得到什么?

塞吉·朗:某种很恶心、乱糟糟的东西.[笑声]不,还会更复杂,它不是一个曲面.如果这个群不是离散的,首先你得问问它是不是闭的并且可看做部分适合条件的某种东西;如果是的,那么,维数会减少.②但是,如果群是离散的,则

---

① W.P.Thurston,"Three dimensional manifolds, Kleinian groups, and hyperbolic geometry", Bull. Amer. Math. Soc., Vol.63(1982).

② 这话有些难懂,我们用一个例子来说明.在 R^2 上,设 τ 是垂直平移一个单位,σ_α 是水平平移一段距离 α,若取 Γ 为所有 σ_α,其中 α 是任意实数,那么,Γ 是闭的但非离散子群,且有 $\Gamma \setminus R^2 \sim R$,故作同一使维数减少 1,若取 Γ_1 为所有 $\tau^i\sigma_\alpha$,其中 i 是整数(可正可负),而 α 是实数,Γ_1 也是闭的但非离散子群,这时有 $\Gamma_1 \setminus R^2 \sim \Gamma_2 \setminus R \sim S^1$.其中 Γ_2 是所有 τ^i 组成的群,而作同一也使维数减少 1,而所得空间 S^1 是紧致空间,这就是部分满足条件的含义.—— 译者注



被你同一的两个点之间有大量的空间,所以作同一时维数不会改变.如果群不是离散的,也可以给你某种可怕的东西——啊,不一定可怕,但有时维数会下降.好吧,你查查书就能弄明白什么事会发生.

男士:刚才,你用双曲几何即双曲距离的概念定义了流形.这是不是只为了方便,或者瑟斯顿理论是否依赖于邻域、开集、闭集这些比双曲距离更一般的概念?

塞吉·朗:当你说邻域、开集、闭集时,你正好是讨论一种没有距离的几何.我列了一张表包含8种几何:

R^3,S^3 和 $S^2 \times R$ 上有普通距离;

$H^2 \times R$ 在 H^2 上有双曲距离,在 R 上有普通距离;

H^3 上有双曲距离;

而其余3种情形有更复杂的距离,也不是双曲距离①.

男士:那么章鱼的拓扑性质应该与这种距离有密切联系,但初看起来它应与此无关?

塞吉·朗:这是一个精妙的注解.瑟斯顿的发现正是两种不同概念之间的联系.这正是他的理论引起广泛兴趣的原因.

高中学生:是的,但你应该对双曲平面上的离散群进行分类,否则你得不出结果.

塞吉·朗:[笑且显出非常高兴的样子]他简直是绝对正确,你叫什么名字?

学生:保尔.

塞吉·朗:实际上,我们在这里做了什么事?我们把一种不知道的东西化成另一种知道的或者不知道的东西.

在数学史上,在某些情况下,我们对离散群知之甚多,而在另一些情况下则不然.我们已经知道得很多,但仍然有很多依然神秘.在19世纪和20世纪中有很多人在研究它们.过去的30年中,关于这些群的研究有重大进展.我们已经对某些群了如指掌,类似地,人们对某些三维流形也知之甚多,而对另外一些却全然不知.所以我的回答是,把对流形的研究化为距离几何对离散群作商的方法,是向前进了一大步.因此我的回答是相对的.

---

① 这里我列出其余三种几何以供数学家参考.其一是 $\widetilde{PSL_2(R)}$,其中 ~ 表示万有覆盖空间.另两种涉及的群是矩阵群.一个是由矩阵

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}$$

组成,叫海森伯格(Heisenberg)群.

另一种由矩阵

$$\begin{bmatrix} a & 1 & b \\ 0 & a^{-1} & c \\ 0 & 0 & 1 \end{bmatrix},a\text{ 为正实数}$$

组成.它们作用的空间等价于 R^3,但使用的距离不同于欧氏距离.

你看,在数学上,经常发生这样的事:有两件我们完全不懂的事情,而我们去证明它们一个等价于另一个.这并非意味着毫无进展,因为问题已经减少了一半.[笑声]但现在发生的事不完全是这样.人们对三维流形有某些了解,也对离散群有所了解.某种意义上说,它们是互补的.把它们放到一起,瑟斯顿能对二者都有了解.

这不是说我本人懂得离散群的分类,这不是我擅长的数学.可以学,但我研究的是别的东西.我知道一些例子,如果你想听,我也可以讲一些,但我对大部分内容知之不多.数学上可做的事很多.当你需要某方面知识时,你可以找一位朋友给你讲,就像我请诺伊曼给我讲瑟斯顿理论那样.

男士:让我们回到二维情形.把一个正方形作同一,你得到环面.但球面怎么办呢?

塞吉·朗:S^2 除外,你不能用同一的办法得出来,至少不能用我这里讲的办法.同样,S^3 也是例外.那是庞加莱猜想.证明紧致、无边界、没有洞的三维流形只能等价于 S^3.这个猜测仍孤傲地站在这套理论的外边.与 S^3 相关联的困难不同于与章鱼相关联的困难.庞加莱猜想是不可约的.①

男士:那低维的呢,如 S^2?

塞吉·朗:至于 S^2,没问题,19 世纪人们就知道答案了,没有洞、没有边界、可定向的二维曲面等价于 S^2.

男士:它可以从平面的表示得出 ——

塞吉·朗:不 …… 嗯 …… 什么样的表示?

男士:用圆盘 H^2.

塞吉·朗:再用离散群?不,有一个定理说不.如果你取一个圆盘,使用庞加莱 - 罗巴切夫斯基几何,再取一个离散运动群,然后作同一,你永远得不到 S^2.这是一个定理.你是数学家吗?

男士:不是.

塞吉·朗:啊,很清楚,数学家应该知道答案.[笑声]啊,不,不!不开玩笑,这个问题很切题,你的良好反应很引人注目.

女士:但庞加莱描述了两种几何,至少我觉得是.

塞吉·朗:好,我们又回到刚才那位男士的问题.他说可以给同一个空间用多种方法赋予距离.在双曲平面上,不只是有我已说到的距离,就是距离变化速率为 $1/(1-r^2)$ 的距离,还有多种方法可以定义距离:有无穷多种定义方法.对这类距离的研究叫做微分几何.它包括研究定义距离的各种可能方法,以及引入某种等价关系来对距离进行等价分类.但真要讲这些就得开一门微分几何的

① 这里作者用了不可约、不可化简(irreducible)一词的双关语义开了一个玩笑.—— 译者注





课程.你是对的,这个课题在很多方面还是公开的,尚未解决的.

女士:但具体说来,并没有实现——

塞吉·朗:啊,具体地说,某人觉得是具体的东西,另一人会觉得很抽象.它牵涉你的头脑,你的知识结构,你在数学上的天分,你的智力,你的品味,你的感觉.它完全是相对的,没有绝对的具体或者抽象的概念.例如,你昨天或者今天觉得太抽象的东西,明天就可能已变得很具体.

如果我画足够多的章鱼,它们会让你觉得很具体.在一定程度上,这是一个习惯的问题,它依赖于环境.没有绝对的答案.当然,数学家可以做一些旁人不能懂的事,旁人会从心理上说这些东西太抽象,这其实是"我不懂"的另一说法.

女士:它没有现实性.

塞吉·朗:什么"现实性"?

女士:物理上的.

塞吉·朗:啊!物理世界比你想的要广阔得多.首先,你知道,你可以取三个表空间的维数,再加上一个表时间的维数,你已经得到四维空间.如果你跑得非常快,你会看到什么?你会看到章鱼吗?你看到四维空间的东西吗?这已经有物理现实性.你的物理现实性终止于何处?你住在一个什么样的空间中?它是弯曲的吗?它是一个章鱼吗?它是不是某种像 H^3 的东西,或者球,而有另一种度量的球?至于发现是什么样的空间,有什么样的度量,是物理学家的事.应该让物理学家来从各种模式中选择,而数学家已经发现了很多模式,也许还会创造更合适的模式.一般说来,人们相信我们的空间是各向同性的.也许实际情况并非如此.

取一个在空间中游走的点.除开它的空间坐标外,它有时间坐标,还可以有速度、加速度、曲率,这些给出其他参数,其他的数值,其他的维数.取一个在空间运动的电子,同时,它还转动,振动,这又给我一些维数.要给这样的电子构造一个模式是很复杂的,甚至给电子概念一个有意义的模式也是这样.要描述这类振动的东西,要描述基本粒子,你需要其他模式,它们也是从微分几何中出来的,当然还涉及别的东西.物理不在任何特定的地方停下来!它不是我可以在黑板上画出来的物理.对其他物理现象.也许我还需要别的模式,它们可能对你来说太抽象了.

女士:是的,当然.[这女士做一个手势,以表示她明白在物理中可以运用的数学模式也可以从其他理论中来,不管多么抽象多么高深的理论]

塞吉·朗:所以,一个好的物理学家应该不被复杂的模式吓倒,他不是小鸡,他会去找他的模式,也许别人觉得太抽象的模式.发现好模式的物理学家将是胜利者.他会出现在科学史上,就是因为他能从束缚别人的思想桎梏中解放

出来,把别人认为太抽象的东西具体化.换言之,没有极限.唯一的极限是每个人个体的,他的头脑、他的素质、他的品位.

[塞吉·朗停下来舒了一口气]

啊![笑]一次马拉松!

[热烈掌声.经过三个半小时演讲,仍有约100人在场]

好了,再见吧.不是天天有这种事,这是唯一的,能够和像你们这样的听众谈三个半小时.真是唯一的!我真的非常欣赏,我真的高兴极了!





漫谈拓扑学

第十八章

1 拓扑学的对象①

拓扑学——最新的几何学科之一,它的创立与欧拉(Euler)、庞加莱(Poincaré)、弗雷歇(Fréchet)、黎斯(Reisz)和其他学者的名字分不开.拓扑学作为独立的数学分支还不过半个世纪之久.

每一数学分支的发展,基本概念贯穿着这部门的所有知识,并且可以说确定了这门数学分支的面貌.那么,引进怎样的基本概念来创立这门新科学——拓扑学呢?这就是连续性概念.连续性概念在数学分析中并不是基本概念,它从属于其他概念,所以多少不能得到完全发展.而在拓扑学中连续性概念得到了完全的、各方面的发展.

1.1 连续函数与连续映象

在数学分析中,我们遇到各种各样的函数,而较常见的是连续函数.在拓扑学中,函数概念与连续性概念得到进一步的扩展.

当我们说 y 是 x 的函数,记作 $y=f(x)$ 时,这就是说 y 依赖于 x,也就是对 x 的每一个值,规定了 y 的某一确定的值与之对应.函数可以由公式给出(例如 $y=\lg(x^2+1)$),也可以通过

① 本节原载苏联《数学教学》(1957～1961).B·Γ·波尔金斯基,B·A·叶夫来莫维奇著.高国士译.

$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$

某些图形的几何性质来给出(如果圆周的半径是已知的,则弓形的弧长是弓形面积的函数),或以其他任何方式给出.此后我们考察函数 $y = f(x)$,总是指 y 是以某种方式依赖于 x 的.

函数 $y = f(x)$ 可能不是对所有的实数 x 而是仅对一部分实数有意义.例如函数 $y = \lg x$ 仅对 x 的正值有意义;函数 $y = \arcsin(\lg x)$,仅对介于$\frac{1}{10}$与 10 间的数(包括端点)有意义.一般说,每一个函数 $y = f(x)$ 有它的定义域,也就是使函数有意义的所有 x 的值的集合.

给出函数 —— 这就是对某集合 A(定义域)的每一点 x 规定了另一集合 B 的确定的点 $f(x)$ 与之对应.函数的这样一般的理解可以叙述为在集合 A 上给定了在集合 B 中取值的函数,或者说给定了由集合 A 到集合 B 的映象.

当讲到一个集合到另一个集合的映象时,所考察的集合并不只是由实数组成.例如设集合 A 是等边三角形的边上的点所成的集合,集合 B 是这三角形的外接圆周上的点所成的集合(图 18.1).把三角形边上的点映成圆周上的点的中心投影就是集合 A 到集合 B 的一个映象.

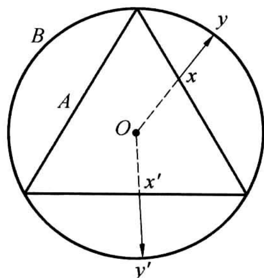


图 18.1

现在转到连续函数概念.函数 $y = f(x)$ 称为在点 x_0 是连续的,如果对于与 x_0 相差很小的 x 的值,函数值 $f(x)$ 与 $f(x_0)$ 也相差很小.说得更精确些,函数 $f(x)$ 在点 x_0 是连续的,如果对任意数 $\varepsilon > 0$,可以找到这样的数 $\delta > 0$,使任何与 x_0 的距离小于 δ 的点 x 所对应的值 $f(x)$ 与 $f(x_0)$ 的距离小于 ε.为了使这定义有意义,必须在集合 A 及集合 B 上规定两点间的距离.在数(实数或复数)的情况,点 ab 的距离可以取作 $|a - b|$.

与点 x_0 的距离小于 δ 的所有点所成集合 U 称为点 x_0 的 δ - 邻域.同样,与点 $f(x_0)$ 的距离小于 ε 的所有点所成的集合 V 称为点 $f(x_0)$ 的 ε - 邻域.这样,数直线上某些点的邻域就是以这点为中心的不大的开线段(不包括端点).如果 A 是球面,x_0 是它的北极,则 x_0 的邻域 U 是球冠(图 18.2).球冠取得越小,点 x_0 的领域 U 越小.对展布在空间的任何图形 A,它的点 x_0 的 δ - 邻域是图形 A 的位于以 x_0 为中心以 δ 为半径的球的内部的部分.

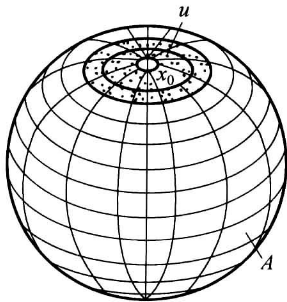


图 18.2

借助于邻域这一术语,连续映象可以定义如下:映象 f 在点 x_0 是连续的,如果对点 $f(x_0)$(在集合 B 中)的任意小的邻域 V 总



能找到点 x_0 的(在集合 A 中)的邻域 U,使 U 中所有点(经映象 f)的象都落在领域 V 的内部.换言之,整个领域 U 映到领域 $\bar{V}$ 内.

容易看到,图 18.3 所描绘的映象在集合 A 的每一点 x_0 是连续的.

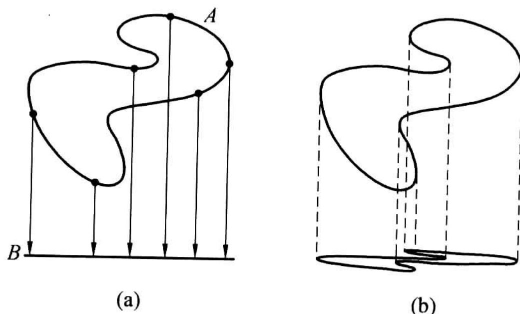


图 18.3

如果集合 A 到集合 B 的映象 f 在集合 A 的每一点 x_0 是连续的,则简称映象 f 是连续的.

在直观上,连续映象把集合 A 中彼此"靠近"的点映成集合 B 中彼此"靠近"的点.也就是经过映象不会发生破裂,不会破坏集合 A 的完整性.注意,可以有集合 A 的不同的点映成集合 B 的同一个点,可能出现"叠合"的情况.

设在平面上有曲线 A 及直线 B,曲线 A 到直线 B 的投影是集合 A 到集合 B 的连续映象(图 18.3(a)),这一映象出现了"叠合",描绘如图 18.3(b) 所示.

考察圆周到图形"∞"的连续映象(图 18.4),它把圆周上的两个不同的点映成图形"∞"的同一个点.

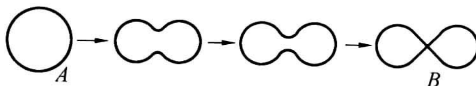


图 18.4

1.2 同胚映象

连续映象不会出现"破裂现象",但可能出现"叠合"情况.有一种非常重要的映象,它既不会出现"破裂",也不会出现"叠合",这种映象称为同胚映象.我们较详细地研究这一概念.

集合 A 到集合 B 的映象 f 称为是双方单值的(即一一对应的),如果集合 B 的每一点 y 正好是由集合 A 的一个点映成.所以双方单值的映象不会把集合 A 的两个不同的点映成集合 B 的同一个点,不会出现"叠合"现象;此外,集合 B 的每一点都为集合 A 的某一点所对应着,也就是集合 A 映成整个集合 B,而不是

集合 B 的一部分. 对于由集合 A 到集合 B 的双方单值映象 f, 可以定义由集合 B 到集合 A 的逆映象: 对集合 B 的每一点 y, 规定集合 A 的这样的一点 x 与之对应, 使得这点 x 经映象 f 的象就是 y. 这逆映象表示为 f^{-1}.

前面图 18.1 所示三角形边界到圆周的中心投影是双方单值映象.

图形 A 到图形 B 的映象 f 称为同胚映象(或同胚的), 如果, 第一, 这映象是双方单值的, 及第二, 又是双方连续的, 也就是不仅是映象 f 是连续的, 而且逆映象 f^{-1} 也是连续的. 上述图 18.1 所示三角形边界到圆周的映象就是同胚映象.

两个图形 A 及 B, 其中一个同胚地映为另一个, 则称图形 A 及 B 是彼此同胚的. 例如, 三角形(或多角形) 的边界与圆周是同胚的.

两个同胚图形在直观上可设想为由某些坚固而有弹性的材料制成, 允许作任何的拉长与弯曲, 但不能撕破和叠合. 如果我们可以通过上述做法把一个图形"安放" 在另一个图形上, 则这两个图形是同胚的. 但是不能设想任何两个同胚的图形总可以在空间通过上述做法把一个图形"安放" 到另一个图形上去. 例如图 18.5(a) 所示的图形(同胚于圆柱体的侧面的带) 及图 18.5(b) 所示的图形(扭转 2π 的带) 是同胚的(为什么?), 但是不能在空间通过拉长、变形(不撕断、黏合) 把一个"安放" 到另一个上去. 这一事实的严格证明是不简单的.



(a) (b)

图 18.5

把图形的同胚概念与全等概念加以比较是有益的. 在几何学里考察称为运动的一些特殊形式的映象, 也就是像刚体那样的位移, 不变动图形各部分相对间的距离(不扭弯、伸长、压缩图形的某些部分). 两个图形, 如果可以借助于运动把一个转移到另一个则称这两个图形是全等的(或合同的), 也就是(从几何学观点看) 这两个图形是一样的, 没有区别的. 在拓扑学里考察的映象, 比运动更一般, 即所谓同胚映象, 这种映象并不保持距离不变, 而是保持图形中点的分布的连续性(也就是不容许撕破与叠合). 从拓扑学的观点看, 两个同胚的图形是一样的, 没有区别的.

下面列举一些彼此同胚的图形.

球面、正方体的面、圆柱体的表面、椭球面等都是同胚的, 但是它们与圆环面(图 18.6(a)) 像汽车轮胎那样的曲面) 不是同胚的. 秤锤(图 18.6(b)) 的曲面和圆环面同胚. 如把俄文字母看做曲线, 则字母 Г, Л, М, П, С 是同胚的, 字母



Е,У,Т,ч,Ш,Ц,Э 也是同胚的,但与前面那些字母不同胚.字母 O 不同胚于任何其他的俄文字母.

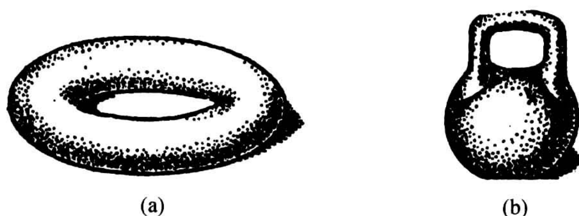


图 18.6

再考察一个同胚的例.设 A' 是以 O 为中心的半圆周,B 是切于半圆周且平行于直径 pq 的直线(图 18.7),由中心 O 的投影把半圆周上的点映成直线上的点.除半圆周的端点 p,q 外,其他的点对应着直线上的确定的点.去掉半圆周的端点 p,q,我们得到图形 A,它同胚地映成(投影) 直线 B.所以,直线同胚于没有端点的半圆周,而半圆周同胚于线段(可以把它拉直).所以直线同胚于没有端点的线段(或称为开线段).

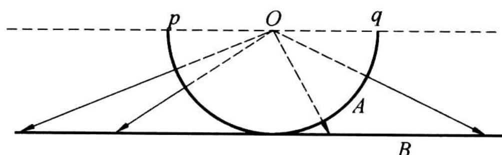


图 18.7

最后的例,我们设 A 是去掉左端点的线段,f 是把这线段连接成圆周 B 的映象(图 18.8).这映象是双方单值的(并没有出现两点的叠合,因作为集合 A 的线段正好只有一个端点).同时容易看出又是连续的.但是,这映象不是同胚的,因为逆映象 f^{-1} 在点 a(对应着线段 A 的右端点) 是不连续的:在集合 B 上在点 a 的左边无限靠近点 a 的点 x,线段 A 上和它对应的点并不靠近线段的右端点,也就是通过映象 f^{-1} 把图形 B"撕破" 了.

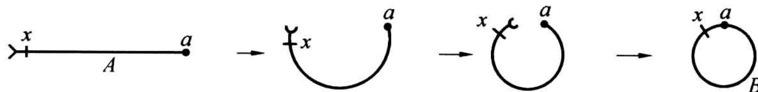


图 18.8

1.3 拓扑不变量

现在已不难阐明拓扑学研究些什么.因为从拓扑学的观点,两个彼此同胚

的图形是算作一样的,所以拓扑学所感兴趣的仅是当一个图形转变为与它同胚的图形时所保持的性质.事实上,两个“拓扑的”一样的图形应该具有“拓扑的”一样的性质.经过同胚映象而不变的图形的性质称为图形的拓扑性质或拓扑不变量.从拓扑学的观点研究图形,就是研究图形的拓扑性质,这就是拓扑学所从事研究的.

三角形的边界具有三个顶点,这性质不是拓扑不变量,因为正方形的边界(与三角形边界同胚)具有四个顶点,而圆周(也和三角形边界同胚)没有顶点.

图形的怎样的性质是拓扑不变量?此外,术语“图形”(这里采用它完全是为了直观)应怎样理解?不阐明这些问题,或多或少不能对拓扑学有一个清楚的概念.我们准备在后面阐明这些重要问题.

2 最简单的拓扑不变量

2.1 拓扑不变量的作用

为了断言两个图形是彼此同胚的,只要实际上得到从一个图形到另一个图形的同胚映象就行了.我们在前面证明三角形的边界同胚于圆周,以及直线与开线段同胚,正是用这种方法.

但是,怎样证明两个图形不是同胚呢?我们找不到一个图形到另一个图形的同胚映象不足以证实这样的同胚映象并不存在.前面我们说过球面不同胚于环面,字母 O 不同胚于其他俄文字母,字母 Г 与 Т 不是同胚的,读者甚至并不怀疑球面与环面不是同胚的.但是,不怀疑是一件事,严格地证明这一事实是另一件事.

为了证明两个图形不是彼此同胚要利用拓扑不变量.比较经常采用的是以数的形式(或其他代数对象的形式)出现的不变量,因为用这样的不变量比较方便.例如,如果我们建立了某种法则,使得相应于每一图形有一个确定的数,并且对应于两个同胚图形的数是相等的.那么,这个数就表示为图形为同胚映象所保持的性质,也就是,这个数是拓扑不变量.如果与两个图形 A 和 B 对应的数是不同的,则它们是不可能同胚的.

近代研究了形形色色的拓扑不变量.这里我们仅考察某些最简单的拓扑不变量.

2.2 连通区的数目

俄文字母 Ы 作为图形,是由两“片”,也就是两个不相联接的部分组成,其他大多数的俄文字母是由联接的一片组成(Й,Ё 除外).组成图形的联接部分



的数目,也就是"片"的数目(称为连通区的数目)是一个拓扑不变量,如果两个图形是同胚的,那么它们是由相同个数的联接部分(连通区)组成.所以字母 Ы 不同胚于字母 О,Л,Ц 等等.图形的连通区的数目是最简单的拓扑不变量.

2.3 分点

在图 18.9(a) 的 ∞ 字形上有这样的点 x,去掉了这一点的尽可能小的邻域后,余下来的是不联接的图形(图 18.9(b))(是由一个以上的连通区组成).具有这种性质的点称为图形的分点.不难看出,∞ 字形上不同于 x 的点都不是分点:对于任何不同于 x 的点 x',总能选取点 x' 的(尽可能小的)邻域(图 18.9(c)),使得去掉这个邻域后,余下的部分是联接的(由一个连通区组成的)图形.

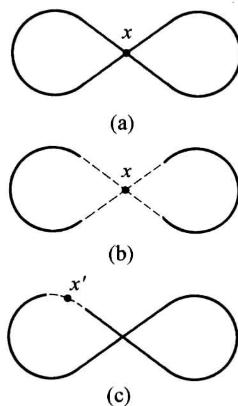


图 18.9

所以 ∞ 字形具有一个分点(图 18.10(a)).圆周没有分点(图 18.10(b)),图 18.10(c) 的图形有两个分点.线段(图 18.10(d))除端点外都是它的分点.图 18.10(e) 的图形除三个端点不是分点外,其他的点都是分点.图 18.10(f) 的图形(同胚于字母 Р,Ђ,Ъ,Ђ) 具有无限个分点,无限个非分点.

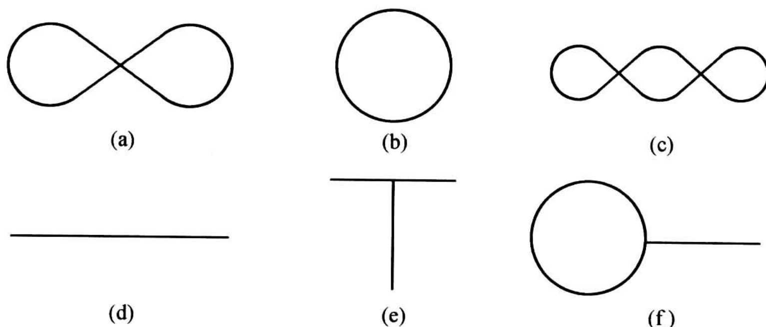


图 18.10

"分点"与"非分点"都是拓扑不变量:如果 x 是图形 A 的分点,f 是图形 A 到图形 B 的同胚映象,则点 $f(x)$ 是图形 B 的分点.所以给定图形的分点的数目是拓扑不变量,非分点的数目也是拓扑不变量.从而可知图 18.10 的图形中,任何两个都不是彼此同胚的.

2.4　点的指数

设图形 A 是由有限个弧组成，x 是这图形的点. 集中在点 x 的图形 A 的弧的数目称为点 x 的指数. 点的指数是拓扑不变量. 图 18.11 所示图形中，点 a 具有指数 1，点 b 具有指数 2，点 c 具有指数 3，点 d 具有指数 4. 指数是 1 的点的数目，指数是 3 的点的数目，指数是 4 的点的数目都是不同的拓扑不变量. 借助于指数概念容易证明字母 Ю 及 Φ 不是同胚的(借助于前面讲过的拓扑不变量不能证明这一事实).

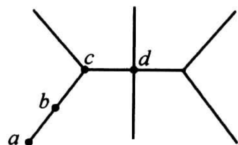


图 18.11

2.5　一笔画成的曲线

有趣的拓扑不变量是图形的"一笔画成"的性质. 由有限条弧组成的图形称为一笔画成的，如果可以连续地通过这图形使每一条弧不经过两次. 图形可以一笔画成这一性质是拓扑不变量. 但是这一拓扑不变量不是新的，可以用点的指数概念来表述：可以证明图形可以一笔画成当且仅当这图形或者不包含具有奇指数的点或者恰好包含两个具有奇指数的点.

2.6　"房屋和井"

图形称为是平的，如果它同胚于平面上某一图形. 图形是平的这一性质是拓扑不变量. 下面两个著名的例涉及这一不变量.

例 1　("房屋和井") 在平面上有六个点：D_1, D_2, D_3(房屋) 及 K_1, K_2, K_3(井)，在平面上能否引出从每一房屋到每一井的小路使其中任意两条不相交? 答案是否定的. 我们可以在平面上画出这样的八条小路，但最后一条(图 18.12 中的 D_3K_1) 在平面上是不能实现的. 所以，由六个点：$D_1, D_2, D_3, K_1, K_2, K_3$ 中的每一 $D_i(i = 1,2,3)$ 向每一 $K_i(i = 1,2,3)$ 所引的 9 条互不相交的弧所构成的图形不是平的. 这图形用 P_1 表示.

例 2　取 5 个点，把其中任意两点用弧连接起来(共 $C_5^2 = 10$(条) 弧)，所得图形用 P_2 表示. 在图 18.13 上画出了 9 条连接弧，第 10 条弧 a_4a_5(虚线) 在平面上不可能画出：它一定与已画出的弧中某一条相交. 可以证明即使这些弧作其他的分布，在平面上引十条不相交的弧是不可能的. 所以图形 P_2 不是平的.

上面所考察两例中的图形是由有限条弧组成，这样的图形在拓扑学中称为有限图(有限网络). 在一般情况，有限图(可简称为图，因为这里不考察非有限的图) 可规定如下：具有有限个点(称为图的顶点)，和连接这些点的互不相交的弧(称为图的棱)；并且图的两个顶点可以用几条棱连接，此外，允许以同一顶

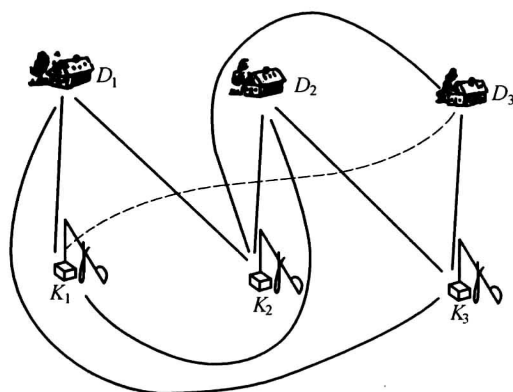


图 18.12

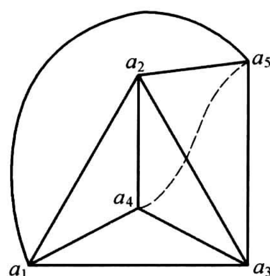


图 18.13

点为起点和终点的棱(称为闭棱).描绘在图 18.14 的图形以及俄文字母都是图的例,前面的 P_1,P_2 也是图.由上所述,图 P_1,P_2 不是平的.有趣的是这两个图是典型的:如果某些图(网络)不是平的,那么它一定包含着(作为一部分)或者图形 P_1,或者图形 P_2.这一结果是邦德里雅金(Понтрягин)于 1927 年证明的,没有发表.后来,这定理的证明为波兰数学家库拉托夫斯基(Kuratowski)独立地得到并发表.

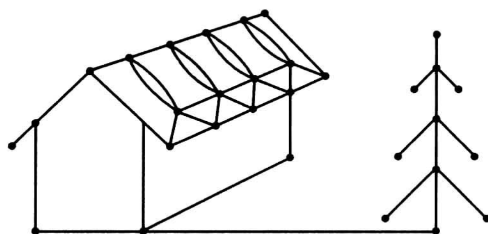


图 18.14

2.7 约当定理

最后,我们还要提到一个拓扑不变量.设 A 是某一平面的图(网络),一般说,这图可以以不同方式分布在平面上.例如图 18.15 所示:这三个图是同胚的,以不同方式分布在平面上.从图 18.15 可以看到这图的三种情况把平面分成相同个数的部分.一般说,平的图(网络)把平面分成部分的数目不依赖于这图在平面上的不同分布.所以这一数目是所考察图(网络)的拓扑不变量.这一事实的证明并不简单;例如,仅有一个顶点及一条棱(闭棱)组成的平的图(平的网络),也就是同胚于圆周的图(图 18.16),分平面为两部分(内部与外部).换言之,平面上的每一简单闭曲线(即同胚于圆周的曲线)分平面为两部分(内部与外部).这就是著名的约当(Jordan)定理.

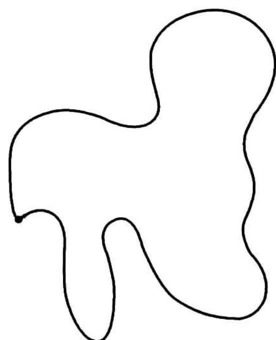


图 18.16

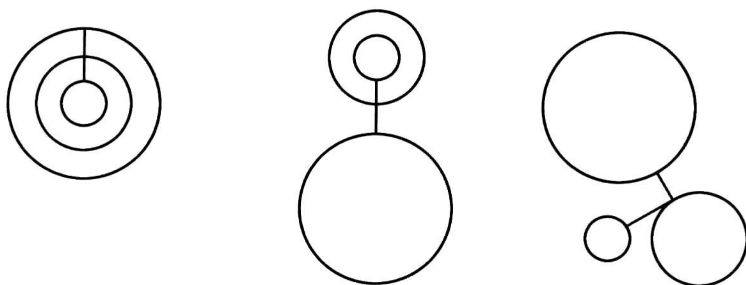


图 18.15

3 曲面的拓扑学

有一个非常有趣而重要的拓扑不变量,首先出现在天才的欧拉的工作中.我们转到这一不变量 —— 欧拉示性数 —— 的叙述.

3.1 欧拉定理

熟知的五种正多面体(正四面体、立方体、正八面体、正十二面体、正二十面体)(图 18.17) 的顶点、棱,面的数目如下:



多面体名称	顶点数	棱数	面数
正四面体	4	6	4
立方体	8	12	6
正八面体	6	12	8
正十二面体	20	30	12
正二十面体	12	30	20

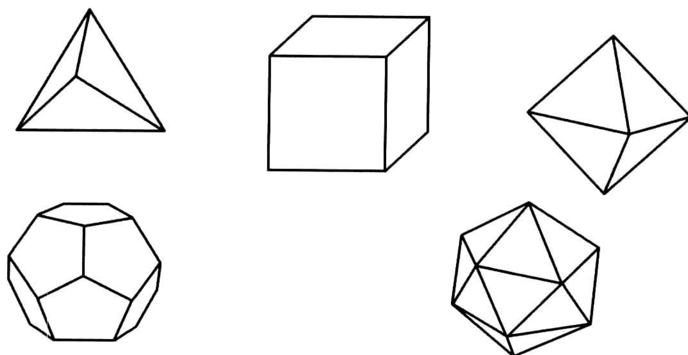


图 18.17

在表上容易看到这些正多面体的顶点数与面数的和比棱数多2.如以 B 表示顶点数,P 表示棱数,Γ 表示面数,则有关系式

$$B - P + \Gamma = 2 \qquad ①$$

关系式①不仅对正多面体成立,对另外一些多面体如棱柱、棱锥等也成立.欧拉首先发现并证明了多面体的这一卓越性质.关系式①现在称为关于多面体的欧拉定理.

现在来说明这定理所表达的内容.首先注意前面所考察的多面体的每一个面都是多角形,也就是圆盘的同胚象.也有这样的多面体,它的面并不都是圆盘的同胚象.例如图 18.18(a) 的多面体的上面的水平面不同胚于圆盘(而同胚于圆环).对这个多面体 $B = 13, P = 20, \Gamma = 10$ 不满足关系式①.这只是因为这个多面体有不同胚于圆盘的面的缘故;如果我们引两条棱把上述水平面分为两部分(图 18.18(b)),那么这两部分都同胚于圆盘,而我们得到:$B = 13, P = 22, \Gamma = 11$ 满足关系式①.所以,我们以后讲到"多面体"时总是指它的每一个面是同胚于圆盘的(在这种情况,每一个面的边界都是一条封闭曲线,而不是像圆环的同胚象那样,它的边界是两条封闭曲线).

上面所考察的每一个多面体的曲面都同胚于球面.事实上,作一球面包围上述每一个正多面体,然后由球心投影把正多面体的曲面投影到球面上就得到

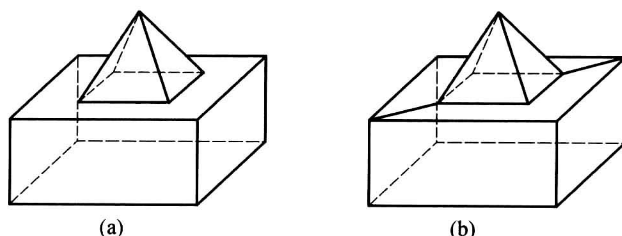


图 18.18

所要求的同胚映象.图 18.18(a),(b) 的多面体的曲面也同胚于球面.多面体的曲面如果不同胚于球面,式 ① 就不能成立.例如图 18.19 的多面体,它的曲面容易看出是同胚于环面的.对于这多面体 $B = 16, P = 32, \Gamma = 16$,所以 $B - P + \Gamma = 0$,式 ① 不成立.这样,在欧拉定理中不明显地以曲面同胚于球面的多面体为前提的(注意任何凸多面体的曲面总是同胚于球面的).因而欧拉定理的明确的形式应是:

对于每一多面体,如果它的曲面同胚于球面,而它的每一个面同胚于圆盘,则关系式 ① 是正确的.

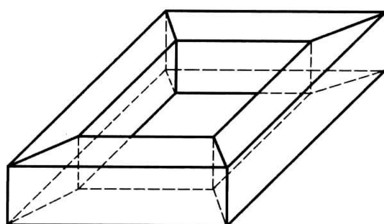
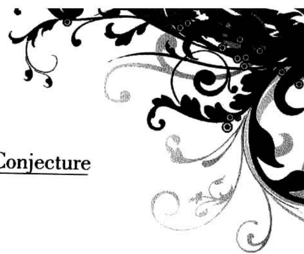


图 18.19

最后,我们给出这定理的纯粹的拓扑形式.首先注意多面体的所有顶点和棱形成一个图(网络,称为所考察多面体的棱网络).如果我们按图的棱把多面体的曲面分成分离的面,也就是同胚于圆盘的片.我们得到下面(比欧拉定理更一般的)断言:

设在球面(或与之同胚的曲面)上构造这样的图(网络)把球面分成同胚于圆盘的片.以 Γ 表示这些片的数目,以 P, B 分别表示图(网络)的棱、顶点的数目,则等式 $B - P + \Gamma = 2$ 是正确的.

欧拉定理的证明不太复杂(可用三、四页篇幅叙述这证明),但是我们在这里不引进这证明.爱好的读者可以在拉杰马赫尔(Г. Радемахер)及盖泼利兹(Р. Теплиш)的书《数与形》(*Числа и Фигуы*)上找到这一证明.



3.2 欧拉示性数

我们考察欧拉定理的某些推论,我们称某些曲面 A 可以剖分成多角形,如果我们可以在这曲面上画出这样的图(网络),把这曲面分成有限个多角形(圆盘的同胚映象).图(网络)的顶点数、棱数分别用 B,P 表示,曲面 A 上由图分成的多角形的个数用 Γ 表示,数 $B-P+\Gamma$ 称为图形 A 的多角形剖分的欧拉示性数.球面(或它的同胚象)可以剖分成多角形.欧拉定理断言,对于同胚于球面的曲面,欧拉示性数不依赖于多角形剖分的选择,而为曲面自身所确定,是曲面的拓扑不变量.

研究另外一些可以剖分成多角形的曲面.如图18.20所示,图形 A 是由球面挖去几个圆形窟窿(称为洞)而得到的,不难作出这图形的多角形剖分.如果对这图形的多角形剖分添增几个圆盘,这些圆盘就是为了得到图形 A 而在球面上割去的,这样得到整个球面的一个多角形剖分.用 B,P,Γ 分别表示图形 A 的顶点数、棱数、面数,而把球面上割去的洞的个数记作 q,则对整个球面的多角形剖分的相应的顶点数、棱数、面数应是 $B,P,\Gamma+q$.由欧拉定理有

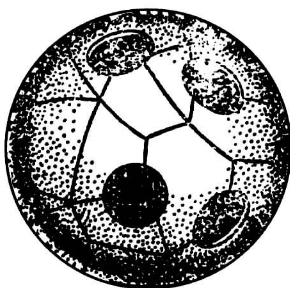


图 18.20

$$B-P+(\Gamma+q)=2$$

故得

$$B-P+\Gamma=2-q \tag{②}$$

因而由球面割去 q 个洞而得到的图形 A 的欧拉示性数是 $2-q$(不管这图形 A 的多角形剖分是怎样的).因此,欧拉示性数是由这图形 A 本身(而不是它的多角形剖分)所确定,也就是拓扑不变量.

可以证明,在一般情况,任意图形的欧拉示性数是拓扑不变量,特别是与图形的多角形剖分无关.欧拉示性数的拓扑不变性具有重大的意义.例如在计算某些图形的欧拉示性数时,可以任意选取这图形的多角形剖分,当然尽量选取最简单的剖分(具有最小的顶点数、面数和棱数).

3.3 黏合

计算某些图形的欧拉示性数.首先我们作如下注记,设 A_1,A_2 是两个图形,各具有同胚于圆周界的边界(图18.21).把这两图形的边界合起来,得到一个新的图形,这一手续称为黏合.同样可以说图形 A_1 的洞被图形 A_2 黏合(或图形 A_2 的洞被图形 A_1 黏合).我们指出,由黏合得到的图形的欧拉示性数等于图形 A_1,A_2 的欧拉示性数的和.事实上,我们可以对图形 A_1,A_2 作这样的多角形剖分,

使它们的边界——圆周——分成相同个数的相等的弧.把它们的边界黏合时,可使顶点与顶点重合,棱与棱重合,这样,我们得到整个被黏合图形的多角形剖分.以 B_1,P_1,Γ_1 表示图形 A_1 的顶点数、棱数、面数,以 B_2,P_2,Γ_2 表示图形 A_2 的相应数目.如果用 m 表示黏合边界上的顶点数(也就是棱数),则黏合后所得图形的顶点数是 B_1+B_2-m(图形 A_1,A_2 顶点数的和是 B_1+B_2,在黏合边界时 m 对顶点重合),棱数是 P_1+P_2-m,面数是 $\Gamma_1+\Gamma_2$,所以被黏合后图形的欧拉示性数是

$$(B_1+B_2-m)-(P_1+P_2-m)+(\Gamma_1+\Gamma_2)=$$
$$(B_1-P_1+\Gamma_1)+(B_2-P_2+\Gamma_2)$$

这证明了我们的断言.

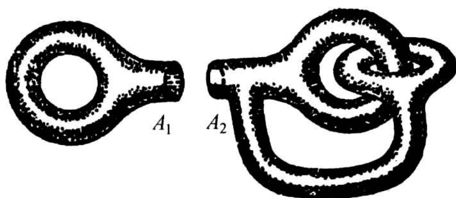


图 18.21

圆盘的欧拉示性数等于1(圆盘同胚于有一个洞的球面,在公式 ② 中让 $q=1$),所以由上述结果得出,用圆盘黏合一个洞时,欧拉示性数增大1.相反地割去一个圆形的洞时,欧拉示性数减少1.由考察过的图18.19的多面体,可知圆环面的欧拉示性数等于零.如果在圆环面上割去一个圆形的洞,则得图18.22的图形,它的欧拉示性数应比圆环面的欧拉示性数减少1从而等于 -1.这一图形称为柄.所以,如果在割去一个圆形洞的图形上,用一个柄黏合这个洞,那么这图形的欧拉示性数减去1.

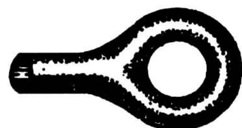


图 18.22

考察已经割去了 p 个圆形洞的球面(如我们所知,这图形的欧拉示性数等于 $2-p$),然后在每一个洞上黏合一个柄,则得到如图18.23的图形,称为有 p 个柄的球面,它的欧拉示性数等于

$$2-2p \qquad ③$$

注意,有一个柄的球面同胚于圆环面(图18.24).有两个柄的球面同胚于由两个柄黏合所得的图形(双环面,图18.25).

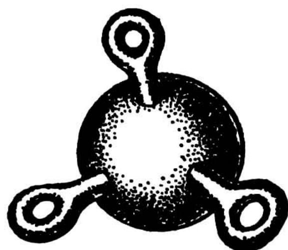
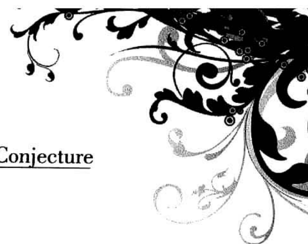


图 18.23



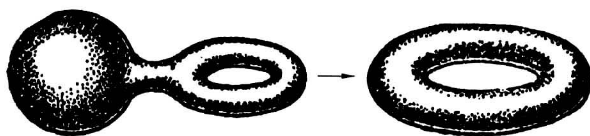


图 18.24

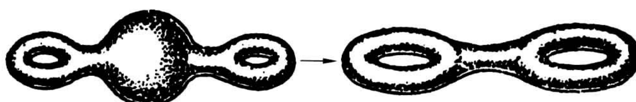


图 18.25

3.4 曲面

球面在任何两点附近是“配置相同的”.为了阐明这情况,我们转向图 18.26 描绘的“华页书” 的图形,考察这图形上所标的点 x,y,z.在这些点的每一点的邻近,图形的配置方式不同.点 x 的邻域同胚于半圆盘并且点 x 处于这邻域的边界上.这时我们称点 x 在图形的边缘上.点 z 的邻域是由三个半圆盘组成,由公共的直径连接着;我们说图形在这里有分支(也就是所考察的图形的三个或更多的叶片邻接着某一曲线).最后,点 y 的邻域同胚于圆盘;这里图形没有边缘,也没有分支.在球面上每一点总具有同胚于圆盘的邻域,也就是球面任何处没有边缘,也没有分支.

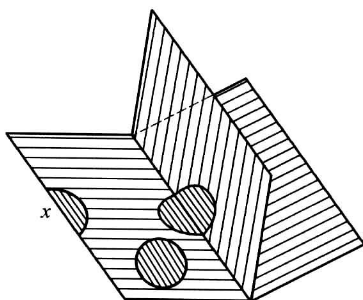


图 18.26

如果图形上的每一点都具有同胚于圆盘的邻域,则称这图形为曲面.曲面没有边缘,没有分支,也就是在它的每一点附近都是拓扑地“配置相同的”.有 p 个柄的球面(图 18.23),特别是球面、圆环面、双环面都是曲面.另外,我们考察具有边缘的曲面,但是它没有分支.圆盘是有边缘的曲面,割去几个洞的球面也是这种曲面.更一般地,在球面上割去 $p+q$ 个洞,而黏合上 p 个柄(图 18.27),

则我们得到有边缘的曲面(有 p 个柄 q 个洞的球面).由前面所述容易推得,这一有边缘的曲面的欧拉示性数等于 $2-2p-q$.

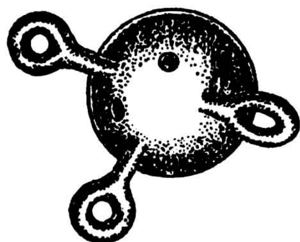


图 18.27

3.5 麦比乌斯带

有边缘的曲面的非常有趣的例首先为德国数学家麦比乌斯(Möbius)及里斯金格(Листинг)在1862 ~ 1865 年的工作所论述.这一曲面称为麦比乌斯带可由下法得到.取易弯曲的长方形带(图18.28(a))卷过一次(图 18.28(b)),然后使两端接近(图 18.28(c))进行黏合,所得到的有边缘的曲面(图 18.28(d)) 称为麦比乌斯带.这曲面的奇特在于它只有一个侧.下面阐明上述含义.

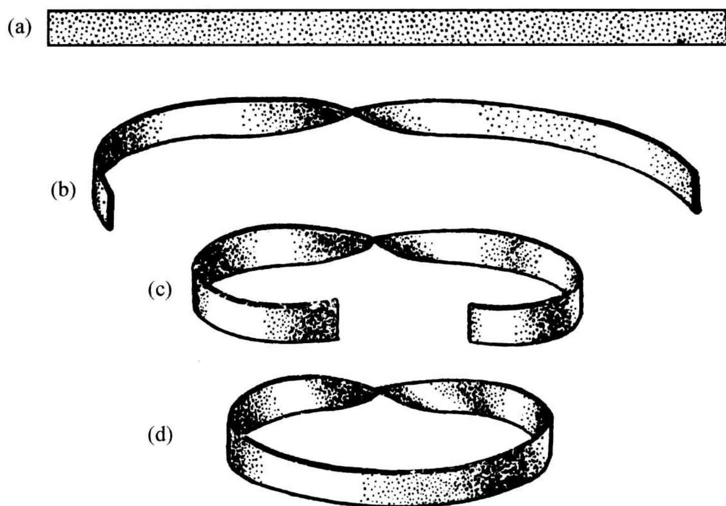


图 18.28

如果长方形的带不卷而黏合两端,则得到有边缘的曲面(图 18.29),它同胚于圆柱体的侧面.这个曲面有两个侧,每一侧可以涂以一种颜色.麦比乌斯带就不能类似地涂上两种颜色:在麦比乌斯带上逐渐地移动毛笔(图 18.30 上毛笔的位置依次标以数字 1,2,3,…),可以到达开始时涂色的地方,但是到了相反的面上.延续地涂色,可把黏合成麦比乌斯带的那条长方形带的两侧都涂上相同颜色.因而麦比乌斯带是单侧曲面.

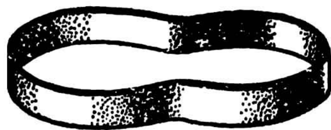


图 18.29

当然,这样借助于"涂色" 的直观描述只能对具有厚度的曲面来说(由某种

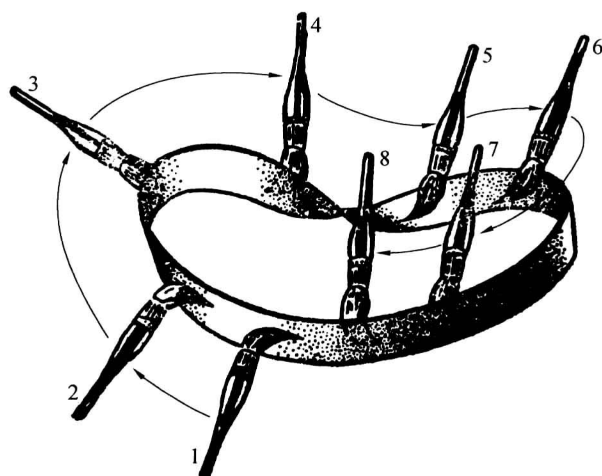


图 18.30

材料组成的曲面),数学上的曲面是没有厚度的.所以我们引进"单侧性"的另外的(数学的)描述.在麦比乌斯带的每一点上可以引垂直于这带的方向相反的两个向量(图 18.31(a)),这两个向量称为麦比乌斯带在点 a 处的法线,选定点 a 处的一个法线,从点 a 开始沿麦比乌斯带移动,法线也随着移动(图 18.31(b)),点 a 绕行整个麦比乌斯带而回到原处时,法线并不是原来的方向而变为相反的方向(图 18.31(c)).因而,在麦比乌斯带上存在着这样的闭路(回路),使当通过这闭路时曲面的法线并不回到原来的方向而是取与原来相反的方向.具有这种回路的曲面称为单侧曲面(建议读者自行验证在球面、圆环面、没有卷过的带上都不存在这种回路).容易理解,单侧曲面的这一数学描述等价于借助"涂色"的描述(用毛笔代替这里的法线).

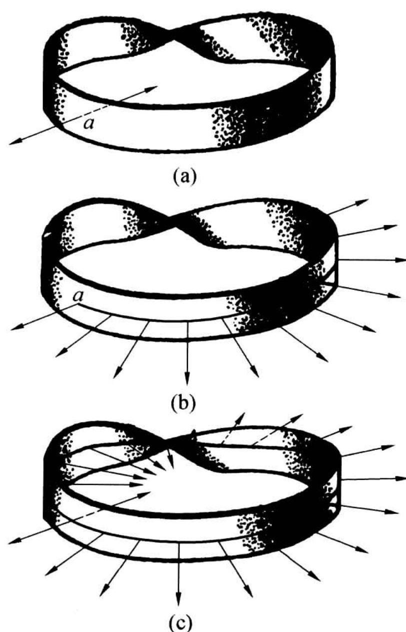


图 18.31

然而,单侧曲面的这样的描述有它自身的不够充分的地方.所给定的曲面

是单侧还是双侧这一问题关系到曲面自身结构的阐明;当我们讲到法线时不仅要研究这曲面,还要研究这曲面在空间是怎样展布的.例如,图 18.5(a) 及图 18.5(b) 的两个图形虽然以不同形式展布在空间,但是它们是彼此同胚的.能否给出单侧曲面的"内蕴的"定义,也就是与图形在空间的展布无关的定义?

由上面叙述,我们不难得到"内蕴的"定义.环绕着点 a(由此引出法线的)画一不大的圆周并规定圆周的方向使从所引法线的顶端观察时这方向是逆时针方向的(图 18.32(a)).如果点 a 移动,法线以及具有方向的圆周也随着移动.当沿着麦比乌斯带环行一周后,点回到它的原来位置,但是法线变为相反的方向.圆周的环绕方向也变为相反的方向(图 18.32(b)).所以,在麦比乌斯带上存在着这样的闭路(回路),当圆周沿着这闭路移动后,圆周的环绕方向改变为相反的方向.这样的回路称为变向的回路.没有变向回路的曲面称为定向曲面或双侧曲面(球面、圆环面、没有卷过的带),具有变向回路的曲面称为不定向曲面或单侧曲面(麦比乌斯带及下文讲的其他曲面).从直观说,曲面的定向性是指可以用小圆周覆盖整个曲面,并选定所有圆周的环绕方向,使附近的圆周具有相同的环绕方向.在不定向曲面上,对充满着曲面的这些圆周,选择如此"协调的"环绕方向是不可能的.

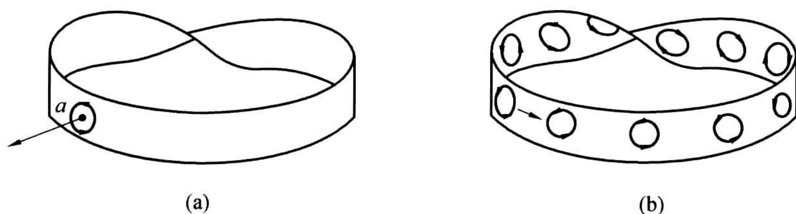


图 18.32

麦比乌斯带的另一有趣的性质是它的边缘同胚于圆周,这一点从图 18.33 容易看到,在这图中右边的麦比乌斯带是用曲线勾画的,这曲线就是它的边缘.

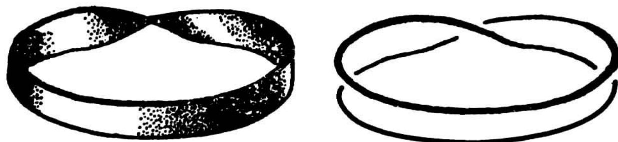


图 18.33

由于麦比乌斯带的边缘同胚于圆周,我们可以设想把麦比乌斯带黏合在某些曲面被切开的圆洞上,使带的边缘对着洞的边缘.为此,我们应将麦比乌斯带扭转、拉长使它的边缘在空间展布为通常的圆周且使麦比乌斯带本身位于这圆周所在平面的一侧;这时黏合可以直接实现.然而把麦比乌斯带在三维空间作这样的展布是不可能的.



为了使人信服这回事,我们可以取两个三角形 amb,cmd,它们相交于线段 mn(图 18.34).另外,再取四个三角形 anc,bnd,amd,bmc 拼合在前面两个相交的三角形上如图 18.35 所示.所得到的图形自身相交于线段 mn 而具有四边形的边缘 $acbd$.我们认为这图形实际上不应具有自身相交的线段 mn,而这自身相交正由于这图形不能"顺利地"在空间展布.我们指出,所构造的图形,如果展开它使自身不相交,是同胚于麦比乌斯带的.为此把这图形沿线段 cm,dm,an,bn 切开,把它们展平形变,然后仍沿这些线段黏合起来.在切开后,图形分为两部分如图 18.36 所示.把它们展平,得到图 18.37 的两平片.这两平片也可以分别形变为图 18.38 的矩形,现在我们进行相反地黏合线段 cm,dm,an,bn.按线段 an,bn 的黏合是把两个矩形合并成图 18.39 的一个矩形.把这矩形卷过来按线段 dm,cm 黏合就得到麦比乌斯带.

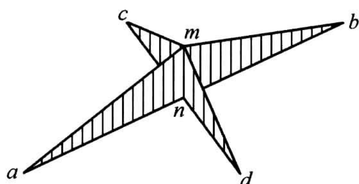


图 18.34

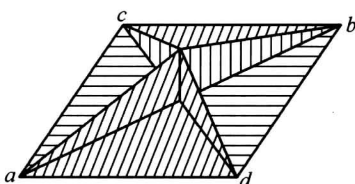
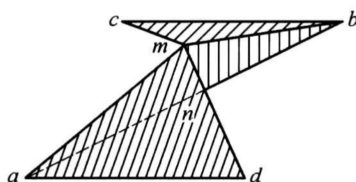


图 18.35



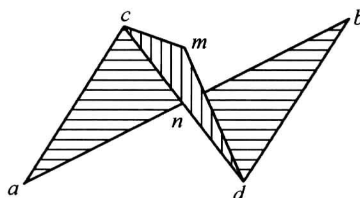
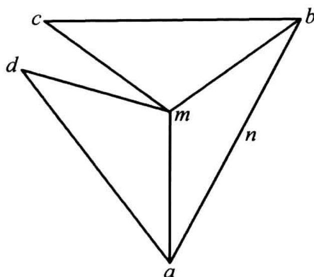


图 18.36



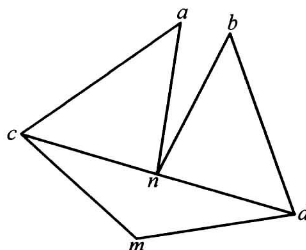


图 18.37

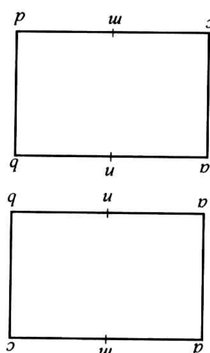


图 18.38

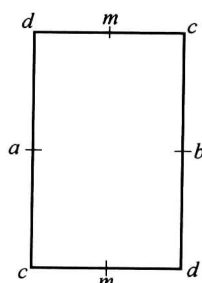


图 18.39

所以图 18.35 所画的是一麦比乌斯带，这时它展布为自身相交的情况.

如果某些曲面被切开一四边形的洞，我们可以把图 18.35 的麦比乌斯带粘贴在洞上，从而得到新的曲面，它在空间展布为自身相交的，这一手续称为“用麦比乌斯带糊洞”.

用麦比乌斯带糊洞常常可用另一方法进行，有时是比较便利的.这一方法是在粘贴时把麦比乌斯带沿它的中线(即图 18.40(a) 矩形的中线) 切开.如果我们把线段 ab 与线段 $a'b'$ 黏合(点 a 与 a' 黏合，点 b 与 b' 黏合)，则由未切开中线的矩形得到麦比乌斯带.为了得到沿中线切开的麦比乌斯带，我们先把矩形沿中线切开(图 18.40(b))，然后像前面一样黏合.

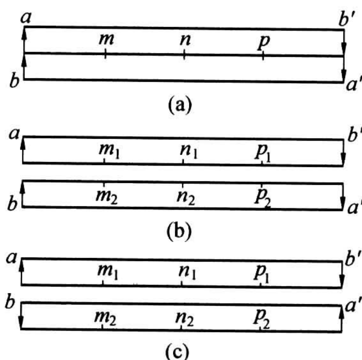


图 18.40

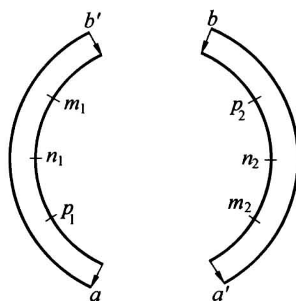


图 18.41

为了黏合，把切开后的矩形的下半个矩形翻过身来(图 18.40(c))，然后把两个半矩形展布如图 18.41，再进行我们所需要的黏合(图 18.42).我们看到按中线切开的麦比乌斯带同胚于圆环.从图 18.40 到图 18.42 沿中线切开后，由点 m, n, p 得到点 $m_1, m_2, n_1, n_2, p_1, p_2$.在图 18.42 上，点 m_1 与点 m_2，点 n_1 与点 n_2，点 p_1 与点 p_2 都分别是直径两端的一对端点(称为对径点).相反地黏合点

m_1与m_2,n_1与n_2,p_1与p_2重又变圆环为麦比乌斯带. 从而得到:把圆环的一个圆周上所有处于直径两端的一对点(对径点)黏合起来得到麦比乌斯带.

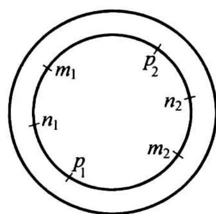


图 18.42

设 l 是某一曲面 A 的圆洞的边缘,在曲面 A 上取一个狭的圆环绕着洞 l,这圆环的外圆周用 l' 表示. 从所考察的曲面 A 上沿 l' 剪去这个圆环(图 18.43),则得到具有较大的洞 l' 的同样的图形(即同胚于原来曲面 A 的图形). 把剪下的圆环的内圆周 l 上所有处于直径两端的一对对点黏合起来,则这圆环形变为麦比乌斯带. 把这麦比乌斯带沿圆周 l' 黏合在曲面上,到此完成了在所考察的图形上(说得确切些是在同胚于曲面 A 的图形上,但这是无关紧要的)用麦比乌斯带糊洞. 但是在曲面 A 上沿圆周 l' 剪开然后再黏合上去这些手续都可以不进行,只要把 l 上所有处于直径两端的一对对点黏合就够了. 所以,在洞的边缘所有处于直径两端的一对对点黏合起来就等价于和麦比乌斯带糊住这洞. 注意,如果我们不允许曲面自身相交,这样的黏合(在三维空间进行)是不能实现的.

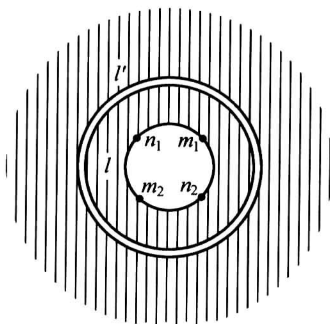


图 18.43

不难相信,麦比乌斯带的欧拉示性数是零,所以用麦比乌斯带糊洞并不改变图形的欧拉示性数. 在球面上割开 q 个洞然后用麦比乌斯带糊没每一个洞,所得曲面的欧拉示性数等于

$$2-q \quad ④$$

3.6 曲面拓扑学的基本定理

现在我们讲述关于曲面的拓扑分类的卓越的定理,这是在上世纪后半叶德国数学家麦比乌斯及法国数学家约当所得到的. 这里考察的仅是闭曲面,也就是这样的曲面,它没有边缘且能剖分成有限个多角形. 例如,平面就不是闭曲面,在平面上画的任何有限图(网络)不能把平面剖分成同胚于圆盘的区域. 曲面的拓扑分类问题包含下述结果,可以给出这样的闭曲面的序列,这些闭曲面两两不同胚,并且任何闭曲面必同胚于序列中某一曲面. 也就是说,可以列举尽所有拓扑不同的闭曲面.

这问题的解答具有如下形式,用 P_0 表示球面,P_k 表示具有 k 个柄的球面,用 N_k 表示这样的曲面,它是把球面切开 k 个洞用麦比乌斯带把这些洞糊没所

得的曲面.我们得到闭曲面的无限集合:

$$\left.\begin{array}{l}P_0, P_1, P_2, \cdots, P_k, \cdots \\ N_1, N_2, \cdots, N_k, \cdots\end{array}\right\} \qquad ⑤$$

我们指出 ⑤ 中的曲面给出了闭曲面的完全拓扑分类,也就是说这里列举尽了闭曲面的所有拓扑不同的类型.

⑤ 中第一行的曲面是两两不同胚的,这是由于它们的欧拉示性数是不同的(由式 ③),同样,第二行的曲面也是两两不同胚的(由式 ④).此外,第一行的曲面都是双侧曲面,第二行的曲面都是单侧曲面(由于用麦比乌斯带糊洞,而每个麦比乌斯带上都有变向回路),所以第一行中的任一曲面不能同胚于第二行中的某一曲面.所以 ⑤ 中的曲面都是拓扑不同的.

至于任一闭曲面总是同胚于 ⑤ 中某一曲面的证明要复杂得多,这里不引进这一证明.(注意,如果我们在球面上割去许多洞,其中的一部分的洞用柄黏合上去,另一部分的洞用麦比乌斯带糊没,这样并不能得到新的曲面.当黏合 p 个柄及 $q(q > 0)$ 个麦比乌斯带时,就相当于用 $2p + q$ 个麦比乌斯带糊洞,这就是曲面 N_{2p+q})

⑤ 中第一行列举的双侧闭曲面可以在三维空间实际地作出,也就是说可以构造这些曲面而不自身相交.第二行列举的单侧闭曲面,如果不允许自身相交就不可能在三维空间构造这些曲面.(在四维空间可以构造任何曲面而不自身相交)

3.7 例

考察某些有趣的曲面的例.

通常的欧几里得平面不是闭曲面.在这平面上添加一些不寻常的点(无穷远点),可以得到闭曲面.这些无穷远点的添加可以如下进行.平面上所引每一直线添加一个无穷远点,对于平行的直线它们具有相同的无穷远点,对于不平行的直线添加不同的无穷远点.这样添加了无穷远点的欧几里得平面称为射影平面.

作一个半球面使和欧几里得平面相切,并且这半球面的直径所在的平面与欧几里得平面平行(图 18.44).我们把半球面的边界点(圆周) 去掉后所得图形称为开半球面.由球心 O 所作的中心投影就是开半球面到整个欧几里得平面的同胚映象(参看前面的图 18.7 及有关叙述).

现在我们进行无穷远点的添加.通过开半球面的切点在平面上引一直线 l,通过点 O 引直线 l' 平行于 l.直线 l,l'"相交于无穷",以至直线 l' 与半球边缘的交点 m_1,m_2(沿着 l') 投影为 l 上同一个无穷远点.所以由具有边缘的半球面到添加了无穷远点的平面(即射影平面) 的映象并不是双方单值的;半球面上



的两个不同的点 m_1, m_2 对应着射影平面上同一个点.为了使这映象双方单值(从而同胚),需要把像 m_1, m_2 那样的两个点黏合成一点.换言之,射影平面同胚于这样的半球面,它的边缘上所有处于直径两端的一对对点相互黏合.也就是在半球面的边缘上黏合一麦比乌斯带,所以射影平面同胚于所得的曲面 N_1,从而射影平面(不同于欧几里得平面)是一单侧曲面.

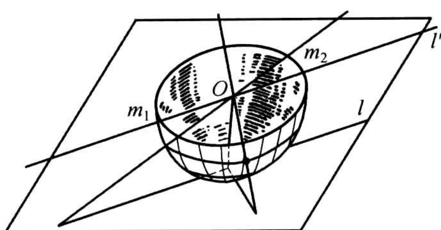


图 18.44

单侧曲面的另一个有趣的例是克莱因(Klein)瓶.考察图 18.45(a)所示有边缘的曲面,图 18.45(b)更清楚地用通过"口"的截面表示这曲面.这曲面具有边缘 l,同胚于圆周,糊没这个圆形的洞得到图 18.45(c)所示的自身相交的闭曲面称为克莱因瓶.这曲面是单侧的:在图 18.45(b)上,从曲面的口的外边的点 a 开始,沿着曲面可以进入"口"的里边.作为单侧闭曲面的克莱因瓶应同胚于闭曲面 N_k 中的某一个.究竟同胚于哪一个呢?下面请读者计算如图 18.45(a)的有边缘的曲面的欧拉示性数,这示性数应等于 -1.由于糊没这个洞后欧拉示性数增加 1,所以克莱因瓶的欧拉示性数等于零,从而克莱因瓶同胚于曲面 N_2.

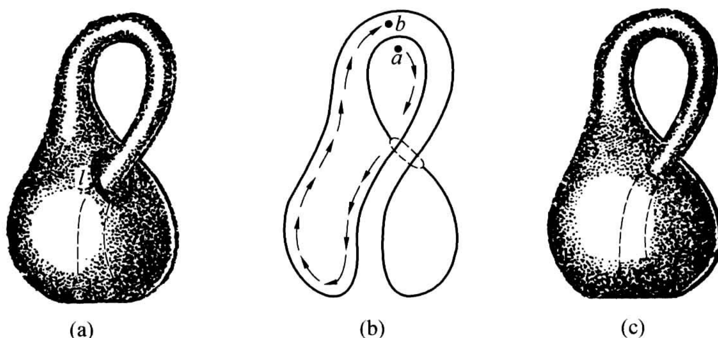


图 18.45

考察双侧曲面一个有趣的例(图 18.46).在正十二面体的一个面上,延长它的边直到相交,我们得到一个正五角星(图 18.47).在相邻的面上构造同样的五角星,这两个五角星具有公共的线段 ad(图 18.48).然而,由于我们要让这两个星形在空间各自展开(实际上是不可能的),我们认为这两个星形仅在线段 ab,cd 处相邻接(bc 不作为邻接的线段),而在线段 bc 处自身相交,be 正是正十二面体的一条棱.对正十二面体的所有面构造同样的五角星,我们得到这样的曲面(图 18.49),它在空间展布为自身相交(自身相交于正十二面体的所有棱处),这曲面是双侧曲面,事实上,可以把每一个五角星涂上两种领色:把向着正十二

面体的中心的一侧(内侧)涂上红色,另一侧(外侧)涂上蓝色.从而,由十二个五角星组成的整个曲面涂上了两种颜色(并且当由一个面转到另一个和它邻接的面时颜色是一致的).

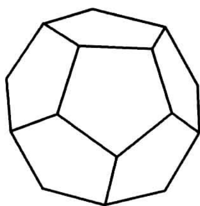


图 18.46

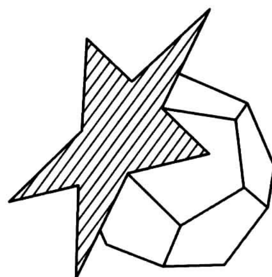


图 18.47

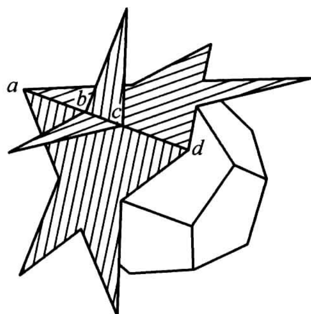


图 18.48

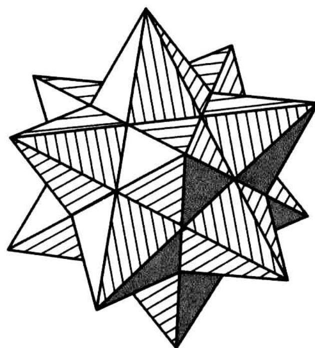


图 18.49

因而,我们所构造的曲面是闭曲面而且是双侧的,所以它应该同胚于闭曲面 P_k 中的某一个,究竟同胚于哪一个呢?我们的曲面有十二个面(五角星).这曲面的棱是由延长正十二面体的棱而得到的(例如延长正十二面体的棱 bc 而得到所构造曲面的棱 ab, cd,注意 bc 是所构造曲面自身相交处,不能作为所构造曲面的棱),所以它的棱的数目是两倍于正十二面体的棱的数目(30),也就是60.最后,这曲面的顶点是32个:原正十二面体的20个顶点都是它的顶点,另外,图上向外凸出的星形体的顶点有十二个(就是正十二面体的面数).所构造曲面的欧拉示性数等于 $32-60+12=-16$.曲面 P_k 的欧拉示性数等于 $2-2k$.解方程 $2-2k=-16$, $k=9$.所以,所构造的曲面同胚于 P_9,也就是具有9个"柄"的球面.



4 抽象几何学

4.1 度量空间与拓扑空间

在初等几何的学习中,可以注意到,关于这些对象(“点”,“线”,……)及其基本关系(点“落在”直线上,点“介于”另外两点间,两线段“重合”,……)的通常直观的说明在严格的逻辑叙述中不起什么作用,而实质上仅是这些关系满足某些准则(公理),在这样的叙述下,所有证明都依赖着这些准则.

从抽象几何学的观点,两个几何系统称为具有“相同的结构”(数学上称为同构),如果在两系统的基本对象间可以建立一一对应,并且在两个系统的基本关系间建立这样的对应,使当一个系统的某些元素满足某关系时,另一系统的对应元素满足所对应的关系,这样的对应称为同构.这样两个系统彼此重复(从所考察的元素间的对应的观点看),彼此成为复制品,好像一国的语言到另一国的语言的逐字逐句的翻译.

例如,系统 A 的元素是某水平面上的点和直线,而另一系统 B 的元素是铅垂的直线和铅垂的平面.两系统的元素间的基本关系用“落在”表示(点落在直线上,直线落在平面上).不难看出,这两系统是同构的:使系统 A 的每一点(或直线)和系统 B 的通过这一点(或直线)的铅垂线(或铅垂平面)相对应.我们得到了系统 A,B 的元素间的一一对应,在这对应下保持着“落在”关系.

抽象几何系统或抽象空间观念明白地表现在“度量空间”概念中,这一概念在20世纪初为法国数学家弗雷歇(Frèchet)引入,在数学中具有很多而重要的应用,并且它的应用甚至超出数学范围.在通常的三维空间(选定了单位度量)对每两点 x,y 确定了它们间的距离$\overline{xy}$(是一正数)满足着熟知的三角形不等式

$$\overline{ab}+\overline{bc}\geqslant\overline{ac}$$

这里 a,b,c 是任意三点(等号当 b 介于a,c 间成立).距离正是两点间的基本关系,它是度量空间概念的基础.

4.2 度量空间

设有任意元素的集合 R(元素称为它的点),对于集合的任意两点 x,y 确定了它们间的距离$\overline{xy}$(或$\overline{yx}$);并设对不同的两点,距离是正数,点到自己的距离为零;此外,对任意三点 x,y,z 满足三角形不等式 $\overline{xy}+\overline{yz}\geqslant\overline{xz}$,在上述情况下,我们称集合 R 形成了度量空间.为便利起见,通常不用$\overline{xy}$ 而用$\rho(x,y)$ 表示距离.我们强调指出距离是点 x,y 的二元函数.

这样就在度量空间规定了任意两点 x,y 间的距离$\rho(x,y)$ 并满足如下条件

(度量空间的公理):

(1)$\rho(x,y)>0$,当 $x\neq y$;$\rho(x,x)=0$.

(2)$\rho(x,y)=\rho(y,x)$(对称公理).

(3)$\rho(x,y)+\rho(y,z)\geqslant\rho(x,z)$(三角形不等式).

抽象集合引进了度量(即确定了距离并满足上述三公理)就转变为度量空间.度量是使集合成为度量空间的唯一关系.如果在同一集合上引进另外的度量则得到另外的度量空间,虽然它们是由相同点集组成.

考察下面的例,在球面上(图 18.50),把任意两点 x,y 间的距离 $\rho(x,y)$ 规定为连接这两点的弦 xy 的长度,这样得到的度量空间用 S 表示.如果把距离 $\rho^*(x,y)$ 规定为球面上连接点 x,y 的最短线段的长度(即连接这两点的两段大圆弧中较短一弧的长度),则得到另外的度量空间 S^*.这样的度量(最短连接线段的长度)对微分几何里所考察的任一曲面都可以这样规定,称为曲面的内蕴度量.最后,在同一集合上还可以规定如下距离 $\rho^{**}(x,y)$:$\rho^{**}(x,y)=1$.当 $x\neq y$;$\rho^{**}(x,x)=0$.这样的度量空间 S^{**} 具有不寻常的性质:球面上的点无论怎样靠近着某一点 x,它和点 x 间的距离 ρ^{**} 并不减小,仅当它和点 x 重合时 ρ^{**} 突然转变为零.

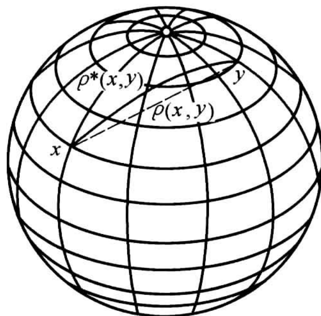


图 18.50

度量空间 S^* 有一有趣的性质:在两点 x,y 间存在着无限个"中点",这是指如下的情况:设 x,y 是球面上处于直径端点的两点(可以看做"两极"),E 是一大圆,它所在的平面垂直于线段 xy(E 可以看做"赤道"),则对"赤道"上的任意点 z,有关系式 $\rho^*(x,z)=\rho^*(z,y)=\frac{1}{2}\rho^*(x,y)$;按照通常的理解,$z$ 可以看做是点 x,y 的中点,整个"赤道"是由这些中点组成.可以指出,空间 S^* 的圆周 E(或任何别的圆周)具有两个中心:x 及 y.

前面的例指出,对我们通常的观念说,度量空间的"几何学"显得非常生疏.

与一般的同构概念相协调,我们应指出,两个度量空间 R 与 R' 称为是同构的(在这情况也称为等距的),如果在它们的点之间能建立一一对应使 R 中任意两点 x,y 间的距离等于 R' 中对应点 x',y' 间的距离.这一对应称为等距对应.例如,在初等几何里,两个图形是等距的,如果一个图形可以通过运动转移为另一图形.此外,在通常的三维空间某一点 O 引出的所有射线的集合,如果规定两射线间的距离为它们间夹角的弧度数,由此形成的度量空间与以 1 为半径的球面在内蕴度量情况下所得的度量空间等距.还可以注意当考察内蕴度量时曲

面的弯曲(是指不改变曲面上任一线段的长度的形变)(图 18.51) 可以作为曲面等距变换的一个直观的例.

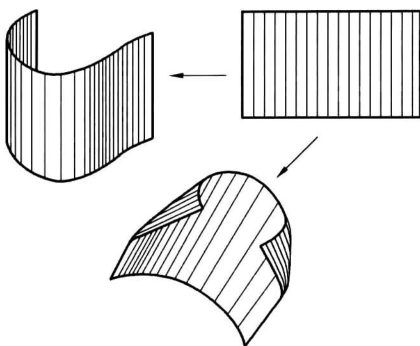


图 18.51

我们还引进些重要的度量空间,所有实数的集合 D 是由距离 $\rho(x,y)=|x-y|$ 形成的度量空间(数直线).所有复数的集合 Z 按同样的距离也形成度量空间(复平面).更一般地,所有 n 个实数组 $x=(x_1,x_2,\cdots,x_n)$ 的集合,规定两点 $x,y=(y_1,y_2,\cdots,y_n)$ 间的距离为

$$\rho(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_n-y_n)^2}$$

从而形成的度量空间 R_n 称为 n 维欧几里得空间.当 $n=1$,得到数直线 R_1,当 $n=2$,得到空间 R_2,它与通常的(欧几里得的)平面及复平面 Z 等距.当 $n=3$,得到空间 R_3,它与通常的(欧几里得的)空间等距.

作为完全不同性质的例,我们考察定义在区间[0,1]上的所有一元连续函数的全体所成的集合 C.对这集合 C 的任意两点(即任意两个连续函数)f 和 g,给以距离

$$\rho(f,g)=\max_{0\leqslant x\leqslant 1}|f(x)-g(x)|$$

所得到的度量空间 C 在近代数学(例如在泛函分析)中起着很大作用.

到此为止,我们讲的是度量空间两点间的距离,现在要讲两集合间的距离的概念.

度量空间 R 的两个集合 A 及 B 称为不靠近,如果存在着这样的正数 α 使对集 A 的任意点 x 及集 B 的任意点 y 都有 $\rho(x,y)\geqslant\alpha$.这些正数 α 中的最大的称为集合 A 与集合 B 间的距离,记作 $\rho(A,B)$.如果具有这样性质的正数 α 不存在,则称集合 A 与集合 B 靠近(或无限靠近),而把这两集合间的距离算作零,在这情况,我们记作 $A\delta B$,读为"A 靠近 B",$A\delta B$ 与 $\rho(A,B)=0$ 所表示的内容是一样的.

图 18.52 ~ 18.55举例说明两集合间的距离概念:用粗线描画的图形 A 及 B

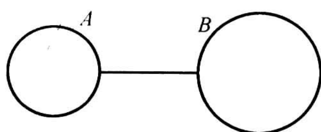


图 18.52

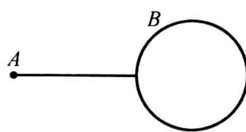


图 18.53

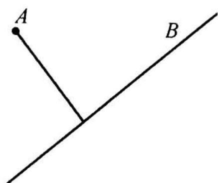


图 18.54

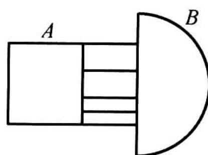


图 18.55

是在同一平面上，用细线描画着由集 A 引向集 B 的最短线段，这最短线段的长度就等于 $\rho(A,B)$. 在图 18.53 与图18.54 中图形 A 是由一点组成. 还应指出，两异面直线间的距离等于它们的公垂线的长度(图 18.56).

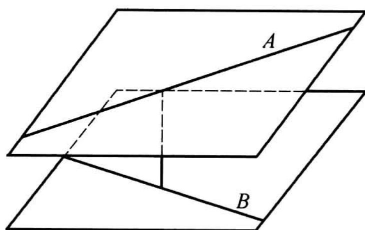


图 18.56

应该注意，在两集合 A,B 中分别找出这样的点来使这两点的距离正好等于 $\rho(A,B)$，这不是完全可能的. 例如同一直线上的没有公共点的两个开线段(除去端点的线段)(图 18.57) 是不可能的情况. 此外，设 A 是某直线，B 是一曲线以 A 为它的渐近线(图 18.58)，则 $\rho(A,B) = 0$，也就是集合 A,B 靠近，虽然它们没有公共点.



图 18.57



图 18.58

4.3 连续性

我们现在指出，在前面所谈的拓扑的概念可以借助于唯一的几何概念——点与集合的靠近来描述. 我们首先观察连续性概念.

设 f 是度量空间 R 到度量空间 R' 的映象. x_0' 是空间 R' 的点，它是空间 R 中的点 x_0 经映象 f 所得的象. 映象 f 称为在点 x_0 是连续的，如果对每一个正数 ε 存在这样的正数 δ，使所有与 x_0 的距离小于 δ 的点经 f 所得的象与 x_0' 的距离小于



ε(也就是点 x_0 的整个 δ - 邻域经 f 所得的象包含在点 $x_0{}'$ 的 ε - 邻域内).

我们证明下述命题:映象 f 在点 x_0 是连续的,当且仅当满足下列条件(*):对每一个靠近点 x_0 的集合 A,它的(经 f 所得的)象 A' 与点 x'_0 靠近(就是从 $A\delta x_0$ 可以导出 $A'\delta x_0{}'$).

事实上,设 f 是在点 x_0 连续的且 $A\delta x_0$,即 $\rho(A,x_0)=0$,则对每一 $\delta>0$ 可在 A 中找到点 x,使 $\rho(x,x_0)<\delta$.设点 x 经映象 f 所得的象为 x',x' 属于 A',且满足 $\rho(x',x'_0)<\varepsilon$($\varepsilon,\delta$ 是按连续性定义选定的),从而对任何 $\varepsilon>0$,在 A' 可找到与点 x'_0 的距离小于 ε 的点,这表示着 $\rho(A',x'_0)=0$,即 $A'\delta x'_0$,所以满足条件(*).

反之,设条件(*)满足,要证 f 在点 x_0 连续.如不然,对某一 $\varepsilon_0>0$,不存在(按连续性定义要求的)正数 δ,即对每一正数 δ(例如,$\delta=\dfrac{1}{n}$,$n=1,2,\cdots$),可以找到这样的点 x(例如,可把对应于 $\delta=\dfrac{1}{n}$ 的点记作 x_n,$n=1,2,\cdots$).使 $\rho(x_n,x_0)<\delta$,但是 $\rho(x'_n,x'_0)\geqslant\varepsilon_0$.把所有的点 $x_n(n=1,2,\cdots)$ 组成的集合记作 A,则有 $\rho(A,x_0)=0$,但是 $\rho(A',x'_0)\geqslant\varepsilon_0$,也即满足关系 $A\delta x_0$,但不满足关系 $A'\delta x'_0$.这与条件(*)矛盾.从而映象 f 在点 x_0 是连续的.

证完.

通过上述命题的证明,我们可以给出连续性的新的定义:由一个度量空间到另一个度量空间的映象 f 是连续的,如果它保持着点与集合的靠近关系.此外两空间是同胚的.如果在它们间能建立双方单值且双方连续的对应,也就是保持着靠近关系的双方单值的对应.这样,同胚是数学中同构概念的特殊情况:它是关于唯一规定的关系 —— 点与集合的靠近 —— 同构的.

拓扑学的定义现在可以叙述如下:

拓扑学 —— 几何学的一个分支,研究着图形的这样的性质,这种性质依赖着仅是唯一的几何概念 —— 点与集合靠近的概念.这种性质称为拓扑的.

4.4 拓扑空间

如上所述,此后对确定数学系统所规定的关系我们自然企图采用关系 $A\delta x$ 而不采用度量作为基础.这就引向了拓扑空间概念.拓扑空间概念是数学家弗雷歇、黎斯、豪斯道夫所引入.苏联数学家乌利松(Урысон)及亚力克山大洛夫对此作了全面的研究.

任何元素(称为"点")的集合 R 称为拓扑空间,如果对它的每个子集合 A 指定这样的点作为与 A 靠近的,并且满足下述条件(称为拓扑空间公理):

1.点 x 靠近点 a(即由一点 a 组成的集合),当且仅当 $x=a$.

2. 点 x 靠近两集合 A,B 的和集 $A\cup B$①(集合 A 的所有点及集合 B 的所有点所成的集合称为集合 A,B 的和集,记作 $A\cup B$),当且仅当 x 至少与 A,B 中一个集合靠近.

3. 如果点 x 和与集合 A 靠近的点所成的集合靠近,则 x 和集合 A 靠近.

与集合 A 靠近的所有点所成的集合称为集合 A 的闭包,记作 $\overline{A}$. 借助于闭包概念,公理 1、2、3 可以用下面等价的公理代替.

Ⅰ. $\overline{a}=a$ 对任意的点 a 成立.

Ⅱ. $\overline{A\cup B}=\overline{A}\cup\overline{B}$.

Ⅲ. $(\overline{\overline{A}})=\overline{A}$(也就是 $\overline{A}$ 的闭包与 $\overline{A}$ 重合).

从这些公理不难导出如下事实:每一集合包含在它的闭包内(即 $A\subset\overline{A}$). 如果 $A\subset B$,则 $\overline{A}\subset\overline{B}$.

前面我们熟悉过点的 ε - 邻域概念. 现在给出更一般的邻域的定义. 拓扑空间 R 的点 a 的邻域是这样的集合 U,它的余集 $R\setminus U$(空间 R 的点凡不属于 U 者所构成的集合称为 U 的余集,记作 $R\setminus U$)与点 a 不靠近,容易看出,点 a 属于它的任何邻域 U.

点 x 称为集合 M 的内点,如果集合 M 是这点的邻域. 集合 G 称为开集合如果它的所有的点都是它的内点②. 开集合概念紧密地联系着闭集合概念:集合 F 称为是闭集合,如果它包含着所有靠近它的点,也就是和它的闭包重合:$\overline{F}=F$. 容易证明:如果集合 G 是开集合,则它的余集 $F=R\setminus G$ 是闭集合. 相反地,如果 F 是闭集合,则 $R\setminus F$ 是开集合.

由此可知,关于开集合的每一命题对应着关于闭集合的某些("对偶的")命题. 相反的论断也成立.

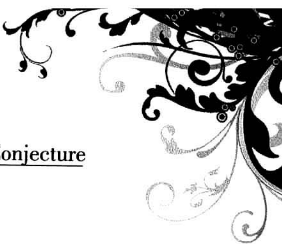
这里是对偶命题的例	
(F)在任何拓扑空间,任意个(可以无限个)闭集合的交集是闭集合. 有限个闭集合的和集是闭集合	(G)在任何拓扑空间,任意个(可以无限个)开集合的和集是开集合. 有限个开集合的交集是开集合

这两个彼此对偶的命题是这样地重要,从其中的每一个可以给出拓扑空间新的定义. 为了使集合 R 作为拓扑空间需要知道它的所有的闭集合(或开集合),并且我们可以按任何方式规定我们所希望的某些集为闭集合(或开集合),只要仅仅满足命题(F)(或命题(G)),我们就可以得到拓扑空间. 换言之,命题

① 与和集相对应的有交集. 集合 A、集合 B 所共有的所有点所成的集合称为集 A,B 的交集,记作 $A\cap B$.

② 与内点概念相对的有边缘点:点 x 称为集合 M 的边缘点,如果它既不是集合 M 的内点又不是它的余集 $R\setminus M$ 的内点. 换言之,点 x 是集合 M 的边缘点,如果它与两个集合 M 及 $R\setminus M$ 都靠近. 也就是 $M\delta x$,$(R\setminus M)\delta x$. 集合 M 的边缘点可以不属于集合 M,例如数直线上的开区间有两个边缘点不属于这开区间. 集合 M 所有边缘点所成的集合称为集合 M 的边缘.





(F)(或(G))等价于公理1、2、3或公理Ⅰ、Ⅱ、Ⅲ.

这样,闭集合(或开集合)可以作为拓扑空间的基本(或原始)的概念,可以借助于闭集合(或开集合)刻画连续的映象.设f是集合R到集合R'的映象,而A'是包含在R'内的集合.R中有这样的点,它(经映象f)的象属于A',所有这种点所成的集合称为集合A'的原象.可以指出,一个拓扑空间到另一拓扑空间的映象是连续的当且仅当闭集合的原象是闭集合(与之对偶的定理是:…… 当且仅当开集合的原象是开集合).

如果在拓扑空间R上给出了实函数(连续函数),对R的每一点规定实数$f(x)$与之对应,也就是空间R到数直线的(连续)映象f.集合F是所有满足$f(x)=0$的R中的点x所成的集合,它是数直线上点0的原象,所以F是闭集合(由公理1,仅由一个点所成集合是闭的).在度量空间,相反的断言也是准确的:对每一闭集合F,可以找到一个连续函数f使F中的点x满足$f(x)=0$.(例如,f可以取作如下函数:$f(x)=\rho(F,x)$)

从上面可以看到,每一个度量空间也是拓扑空间,但是相反的提法是不准确的,存在着这样的拓扑空间,它不可能使成为度量空间.

怎样的拓扑空间可使成为度量空间呢?这一问题显得很重要,这一问题的部分情况的解决是属于苏联数学家乌利松,关于它的工作下面还要叙述.这一问题的完整的表述近来为斯米尔诺夫(Смирнов)及日本数学家长田润一(J. Nagata)所解决.

4.5 连通性

前面提到的由整个一片(连通)组成的集合的性质需要给以严格定义.这一定义属于豪斯道夫.

我们为R中的两集合A,B是拓扑无关的,如果其中每一个不包含与另一个靠近的点.集合M称为是连通的如果它不可能分成两个不空的拓扑无关的集合;也就是由关系式$M=A\cup B, A\neq 0, B\neq 0$可以导得$A,B$是拓扑相关的.可以指出,线段是连通集合.如果在线段中去掉任何一个内点x,那么余下的集合就不连通了,因为在点x的左边及右边部分是两个拓扑无关的集合.

对于连通集合,下面关于连续函数的定理成立.在连通集合M上的连续实函数如果在集M在二点a,b上的值具有相反的符号,则在M在某些点上的值为零.作为推论我们有:在连通集合M上的连续实函数取介于$f(a),f(b)$间的一切值,这里a,b是集M的任意两点.特别是对于定义在区间上的连续函数是正确的.

4.6 一致连续性

度量空间R到度量空间R'的映象f称为是一致连续的,如果对于每一数

$\varepsilon > 0$ 存在这样的数 $\delta > 0$(依赖于 ε) 使 R 中相距小于 δ 的任意两点的象间的距离小于 ε. 注意,一般说 f 的连续性不能导出它的一致连续性. 例如数直线 OX 到数直线 OY 内的映象 $y = x^2$ 在整个数轴上是连续的,但不具有一致连续性. 事实上,对所取的任何 ε 和 δ,可以找到两数 x 和 x',它们间的距离小于 δ 而相应的 $y(= x^2)$ 与 $y'(= x'^2)$ 间的距离大于 ε. 可设

$$x = \frac{\varepsilon}{\delta}, x' = \frac{\varepsilon}{\delta} + \frac{\delta}{2}$$

$$y' - y = (\frac{\varepsilon}{\delta} + \frac{\delta}{2})^2 - (\frac{\varepsilon}{\delta})^2 = \varepsilon + (\frac{\delta}{2})^2 > \varepsilon$$

容易证明,设映象 f 是一致连续的,则它把靠近的集合映成靠近的集合. 事实上,由于一致连续性,对任意小的 $\varepsilon > 0$ 可选取 $\delta > 0$ 使 $\rho(x, y) < \delta$ 时有 $\rho(x', y') < \varepsilon$. 现设 $A\delta B$,即 $\rho(A, B) = 0$,因此对任何 δ,总可找到 A 中的点 x 及 B 中的点 y 使 $\rho(x, y) < \delta$,从而 $\rho(x', y') < \varepsilon$. 设 f 把集 A 及 B 分别映成集 A' 及 B',则 x' 是 A' 中的点,y' 是 B' 中的点. 由于 ε 可以取得任意小,所以 $\rho(A', B') = 0$,也就是 $A'\delta B'$.

相反的论断(需要较复杂的证明)也是正确的:设映象 f 把靠近的集合映成仍靠近的集合,则映象 f 是一致连续的. 这样,一致连续映象可以描述为保持集合间靠近的映象. 我们不能这样想由此似乎可以得出一致连续性是拓扑性质. 要知道拓扑学所刻画的仅是点和集合靠近的关系而并不是集合间的靠近的关系. 每一性质,如果它可以用集合间靠近的术语来叙述的话,则称这种性质为靠近性质,所以一致连续性是靠近性质不是拓扑性质.

两图形称为是等胚的如果在它们的点间可以建立保持集合间靠近的一一对应(也就是双方单值双方一致连续的对应 —— 等胚). 如果两个图形是等胚的,自然它们是同胚的. 相反的论断不是总能成立的. 例如,一端无限伸展的带(图 18.59)、角(图 18.60) 及由曲线 ab、它的渐近线 cd 和线段 ac 所围的平面部分(图 18.61) 都是同胚的,但不是等胚的. 此外,开区间与整个数直线是同胚的但不是等胚的. 旋转抛物面(图 18.62) 与平面是同胚的但不是等胚的. 关于欧几里得平面和罗巴契夫斯基平面也是同样情况.



图 18.59

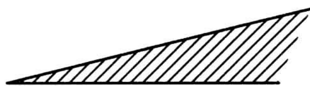


图 18.60

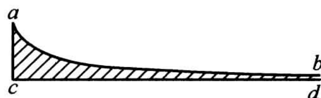


图 18.61

作为例子,我们可证明开区间与数直线不是等胚的.为此,我们对由数直线 R 到(去掉端点的)线段 R'(为方便起见线段的长可以算作1)的任何映象 f,在 R 上找出两个不靠近的集合 A 及 B 而它们的象 $f(A)$ 及 $f(B)$ 是靠近的(当 f 是一一对应时,f 的逆映象把靠近的集合映成了不靠近的集合).在数直线上任取三个整数,因为在长度为1的线段上不可能找出三个点使两两间的距离都大于$\frac{1}{2}$.可在三个整数内选取二个使它们的象间的距离不大于$\frac{1}{2}$.这两个整数记作 m_1 及 n_1.再取不同于 m_1,n_1 的四个整数,因为在长度为1的线段上不可能找出四点使两两间的距离都大于$\frac{1}{3}$.可在这四个整数中选取两个 m_2,n_2 使它们象间的距离不大于$\frac{1}{3}$.用同样的方法取不同于 m_1,n_1,m_2,n_2 的整数 m_3,n_3 使它们象间的距离不大于$\frac{1}{4}$,依次继续下去,设 A 是 $m_1,m_2,m_3,\cdots$ 所成集合,B 是 $n_1,n_2,n_3,\cdots$ 所成集合,显然 A,B 是不靠近的(因为它们的元素是不相同的整数),而 $f(A)$ 与 $f(B)$ 是靠近的.

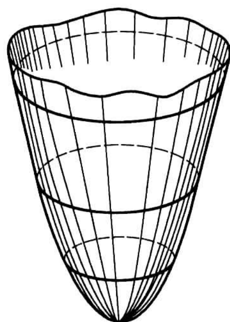


图 18.62

4.7 近性空间

正如拓扑空间作为由唯一规定的关系——点与集合靠近的关系——所定义的数学系统,我们可以考察这样的数学系统,这里所规定的是集合与集合靠近的关系.这就导向近性空间概念,这是由叶夫来莫维奇引进的.

集合 R 称为是近性空间如果对它的任意两个不空的子集可以指出它们作为靠近的或不靠近的并且满足如下公理(近性空间公理):

1.点 a 与点 b 靠近($a\delta b$),当且仅当 $a = b$.

2.集合 C 与集合 $A \cup B$ 靠近,当且仅当 C 与集合 A,B 中至少一个集合靠近.

3.如果集合 A,B 不靠近,则存在着这样的集合 A_1,B_1 使 $A_1 \cup B_1 = R$,且 A_1 与 B 不靠近及 B_1 与 A 不靠近.

两个近性空间称为是等胚的,如果它们按靠近关系同构,也就是如果它们间能建立保持靠近关系的一一对应.

每一度量空间都是近性空间,因为在度量空间可以定义集合间的靠近的关系.每一近性空间又都是拓扑空间,如果把点 a 与集合 B 的靠近理解为仅由一

点 a 组成的集合与集合 B 的靠近.

以前遗留下来的不明确的图形概念(这种在拓扑学里所考察的图形是度量空间、近性空间、拓扑空间以及这些空间中的集合(常常只是闭集或开集))这里把它们精确化了.然而在下面,为了简单和直观起见我们保留着某些不定义的(关于图形的)术语.

5 关于曲线概念

在这里,从最简单的曲线 —— 直线段 —— 出发,试图标出被明显地描述的,自然地称为曲线的一类图形.这可由不同方式去进行.

5.1 简单弧

欧几里得定义曲线为"没有宽度的长".当然这不是严格的数学定义,而是自然地称为曲线的一类图形的直观的描述,然而,下面可以看到,这一直观描述未必成功.

自然地,曲线概念可作为拓扑的概念.当所有图形同胚于我们所认为曲线的图形中的任何一个时,我们也应该称它们为曲线.因此,同胚于直线段的任何图形 —— 这种图形称为简单弧 —— 可以作为曲线.我们试图考察简单弧概念作为曲线概念的第一次推广.圆周上的一段弧,俄文字母 М,И 都是简单弧,而字母 Φ 中不是简单弧.

可是,如下面所指出者,事情并不简单.任何简单弧是否可以作为"没有宽度的长"?也就是这种图形是否不占有任何面积?答案是否定的:存在着平面上的简单弧占有面积!可用下面的方法构造这种曲线的例.

取面积为 1 的正方形(图 18.63(a)),在它中间去掉一个十字形(图 18.63(b)),取适当的宽度使这十字形的面积为$\frac{1}{4}$(容易计算出宽度可取为 $1-\frac{\sqrt{3}}{2}$).这时留下来四个小正方形称为所构造的一级正方形.在每一个一级小正方形内再去掉一个十字形,取适当的宽度使这四个小十字形的面积的和为$\frac{1}{8}$.这时留下 16 个(二级)小正方形(图 18.63(c)).再对一个二级小正方形进行上述手续,使去掉的十六个小十字形的面积和为$\frac{1}{16}$.这时留下 64 个(三级)小正方形(图 18.63(d)).继续进行这样的手续无限回,留下的正方形越来越小,所得(趋极限过程中的)图形的面积也越来越小,看来将趋近零.但事实并不这样:"极限"图形具有面积$\frac{1}{2}$.



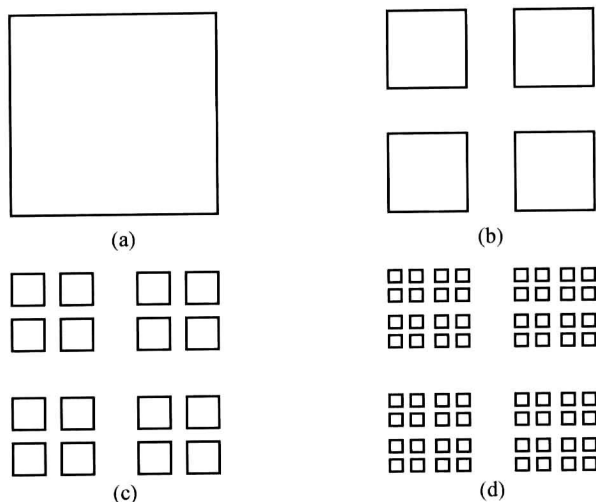


图 18.63

事实上,第一回在正方形内去掉部分的面积是$\frac{1}{4}$,后来依次为:$\frac{1}{8}$,$\frac{1}{16}$,…,所以留下图形的面积依次为$(1-\frac{1}{4})$,$1-(\frac{1}{4}+\frac{1}{8})$,$1-(\frac{1}{4}+\frac{1}{8}+\frac{1}{16})$.继续这样推理,我们相信在极限情况所得图形具有面积$1-(\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots)$.因为括弧内的无限递减等比级数的量是$\frac{1}{2}$,故知"极限" 图形的面积为$\frac{1}{2}$,这图形是很惊人的:它"分散" 为个别的点(因为留下的正方形越来越小并且彼此孤立),但是具有正的面积,这一"极限" 图形用 A 表示.

上述图形 A 是被描述为经过无限回构造手续而得到的"极限" 图形.下面是图形 A 的严格定义:点 x 属于图形A,如果在任一回构造手续时不被去掉,我们也可用另外方式定义(借助于图形的交集):用 A 表示经过n 回构造手续后所留下的图形.则 A 是所有这些图形 A_n 的交集(可以表示为 $A=A_1\cap A_2\cap A_3\cap\cdots$,或 $A=\bigcap_{n=1}^{\infty}A_n$).

为了找出具有正的面积"曲线",我们现在构造简单弧通过集合 A 的所有的点.这简单弧包含整个图形 A,具有不小于$\frac{1}{2}$ 的"面积".

下面转入这一曲线的构造(图 18.64(a) ~ (d)).构造像俄文字母 П 那样的条形区域(图 18.64(a))使包含着前面图形 A 构造过程中的(四个)一级正方形(图 18.63(b)).构造图 18.64(b) 的(弯曲的)条形区域,它包含着图 18.63(c) 的(十六个) 二级正方形.构造图 18.64(c) 的(更弯曲的)条形区域包含着图 18.63(d) 的(六十四个)三级正方形.下一步是构造图 18.64(d) 的条形区域(越

来越弯曲了),依次继续下去.

第 n 回构造的弯曲的条形区域记作B_n,它包含在所有前面构造的条形区域内($B_n \subset B_{n-1} \subset \cdots \subset B_2 \subset B_1$)且包含着图形 A_n(从而包含着图形 A).区域 $B_1, B_2, \cdots, B_n, \cdots$ 的“极限”图形记作 B. B 可以正确地定义为所有 B_n($n = 1, 2, \cdots$)的交集,即 $B = \bigcap_{n=1}^{\infty} B_n$.显然,图形 B 包含着图形 A.考察这一继续无限回的构造过程(图 18.64),我们直观地相信极限图形 B 是非常“迂回曲折”的曲线(简单弧).它具有正的面积①从而未必能称它为“没有宽度的长”.

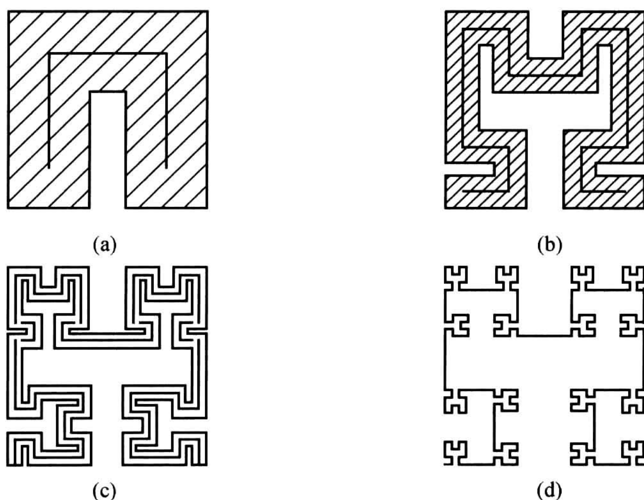


图 18.64

5.2 **道路**

常常给出曲线的另一直观的描述:“曲线是运动着的点的轨迹”.这一描述引向“曲线”的另一观点 —— 道路概念.

当对曲线作这样描述时,占首要地位的是出现了完全不同的概念:曲线并不考察为图形(空间的点集合)而是作为点的有序集合,重要的是运动着的点以怎样的程序通过不同的位置.例如,点以两种不同方式跑过图形 Φ 如图18.65所示(黑线表示已跑过的道路,虚线表示将跑去的道路),在这两种情况运动着的点的所有位置的集合是一样的,但在情况 18.65(a) 及 18.65(b) 以不同的程

---

① 必须指出,我们并没有引进图形 B 的性质的论证.在几何学里的面积概念仅对简单的图形加以定义,我们事先并不清楚,是否可以对如此复杂的图形 A 及 B 谈它们的“面积”.为此,必须把面积概念精确化并加以扩充.此外,在图形内去掉有限部分后所余下部分的面积可以从原来图形的面积中减去这有限部分的面积的和而得到的.这一性质是否对在图形内去掉无限部分的情况也保持呢?最后,可以证明图形 B 实际上是简单弧.由于它包含着集合 A 可以推出它的面积不小于图形 A 的面积.综上所述,要想完备地构造这例(以及下面要讲的例)必须引进更复杂的讨论.

序跑过这一集合.换言之,运动着的点的道路是不同的,虽然在两种情况,点跑过同样的集合.

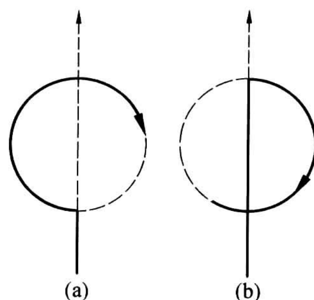


图 18.65

这里引向道路概念的精确定义.设在某拓扑空间或度量空间 R 上(例如通常的平面或三维空间)点从时刻 $t=0$ 到时刻 $t=1$ 连续地"运动".这表示着对每一时刻 $t(0\leqslant t\leqslant 1)$ 点的运动位置 x_t 是已知的,并且随着 t 的变动,点 x_t 连续地变动.换言之,对线段 $0\leqslant t\leqslant 1$ 的每一个点 t 规定着空间 R 的点 x_t 与之对应,也就是线段[0,1]到空间 R 内的映象,并且这映象是连续的,因为点 x_t 随着变量 t 连续地变动.这一映象就是道路,因而,我们可以有下述定义:(用 I 表示线段[0,1]).

线段 I 到空间 R 内的每一个连续映象称为空间 R 中的道路.

道路概念在拓扑学中起着重要作用,我们将在第七章里详细地考察它.在这里我们想把道路概念作为曲线概念的一个重要推广来考察.作为道路来考察曲线有它的优越性:例如可以考察自身相交的曲线(图 18.66(a)).但是同样的曲线可以表示为不同形式的道路(图 18.66(b)、(c)、(d)).

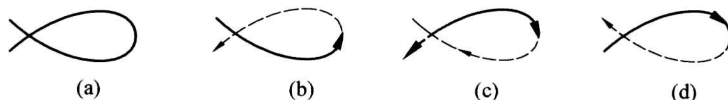


图 18.66

不能想象,道路概念是完全直观而简单的.首先,我们知道任何简单弧本身都是道路(因为简单弧是线段的同胚映象,从而是连续映象).所以,如前面指出的"动点的轨迹"(道路)可以通过具有面积的集合的所有点,也就是道路可以填满面积.不仅如此,下面构造这样的道路,它填满整个正方形,也就是通过正方形的每一个点(换言之,存在着线段到正方形上的连续映象).这种类型的道路称为皮亚诺(Peano)曲线.

为了得到皮亚诺曲线,我们的构造方法完全类似于图 18.64 情况,只是这里的条形区域更宽些使每一个条形区域能填满正方形.这些条形区域见图 18.68,它们是这样得到的:用平行于正方形边的线段把它分成 4,16,64,256,…个相等的小正方形(图 18.67).然后沿着构成小正方形的某些边切开(图 18.68),得到填满正方形的弯曲的条形区域("迷宫","复杂的曲径").图 18.68 上的虚线表示条形区域的中线.这些中线和图 18.64 的条形区域的中线具有同

样的性质.构造过程无限地继续下去,所得到的条形区域的中线越来越弯曲.这些中线的极限给出了填满整个正方形的道路.

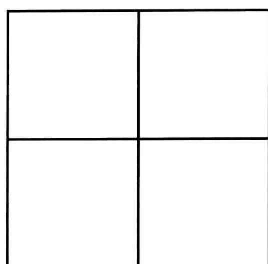
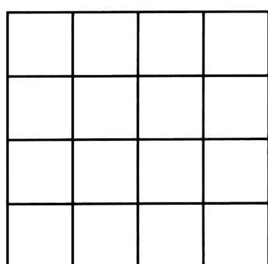
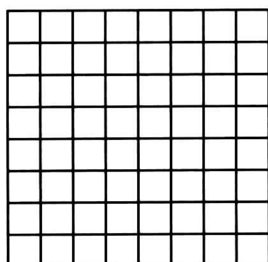
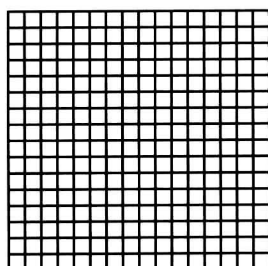





图 18.67

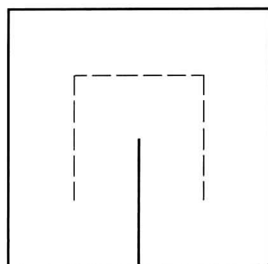
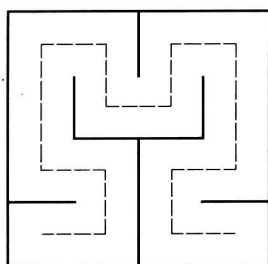
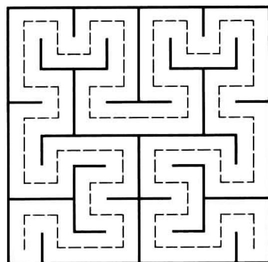
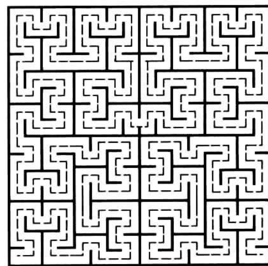
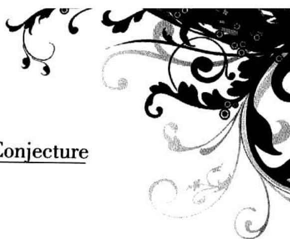





图 18.68

当然,所有这些推理带有描述性质,需要作严格的论证.还要注意所得到的"曲线"(皮亚诺曲线,也就是填满正方形的曲线)不是简单弧:它有无限个点是由几个点"黏合"而成的.也就是在这正方形上有无限个这样的点,道路不止一次地通过它们;换言之,这一由线段到正方形上的连续映象不是双方单值的.





5.3 康托曲线

欧几里得也曾把曲线描述为“曲面的边界”.这样的描述是不成功的.曲面概念比曲线概念复杂,借助于较复杂的概念去定义较简单的概念显然是不合理的.可是,限制在平面曲线情况,我们可以试图以平面上的开集合作为“曲面”,也就是称平面上的开集合的边界为平面曲线.这样的定义实际上是成功的.然而,从下面的例可以看到这一曲线的定义包含着许多出乎意料的情况.

我们习惯地认为曲线的每一部分“分开”平面,也就是平面以“两个方向”与曲线邻接.例如圆周分平面为两部分(内部与外部),而其中的每一部分都以圆周为它的“边界”(图 18.69),可以说圆周是这两部分的公共边界.一般说,曲线 A 称为两个开集 G_1, G_2 的公共边界如果满足下列两条件:

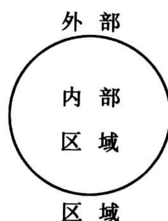


图 18.69

(1) 曲线 A 分开区域 G_1, G_2,就是说联结 G_1 的点与 G_2 的点的任何曲线(简单弧)不可能与 A 不相交.

(2) 两个区域 G_1, G_2 全部地邻接曲线 A,也就是曲线 A 的任何点的任意靠近处总有 G_1 及 G_2 的点.

直观上似乎是明显的,任何曲线不能作为平面上多于两个区域的边界,但这并不正确.下面可以看到:存在着这样的曲线是平面上三个(或更多个)区域的公共边界.这样的曲线分平面为三部分,这三部分与这曲线全部地邻接.这类曲线的构造与日本数学家和田(Wada)的名字分不开,我们引进的描述(仅是描述,不是严格的数学构造!)类似于 1917 年日本数学家米山(Yoneyama)所给出的.设有一四周环海的陆地,上面有两个湖:暖水湖与冷水湖(图 18.70).我们要开水道把两个湖里的水及海里的水引到陆地上.第一天从暖水湖引出水道 W_1,使不与冷水湖及海沟通,并使陆地上的每一点到暖水湖的距离不大于 1(单位长度),第二天从冷水湖引出水道 C_1 使不与暖水湖和水道 W_1 及海沟通,并使陆地上的每一点到冷水的距离不大于 1.第三天从海引出水道 S_1 使不与暖水湖和水道 W_1、冷水湖和水道 C_1 沟通,并使陆地上的每一点到海水的距离都不大于 1.所以三天后,陆地上的每一点到冷水、暖水、海水的距离都不大于 1.以后的三天再从 W_1, C_1, S_1 引出水道使陆地上的每一点到冷水、暖水、海水的距离都不大于$\frac{1}{2}$,以后三天作同样处理,使相应的距离不大于$\frac{1}{4}$.继续下去,水道网愈来愈密,最后(极限情况)得到“到处稠密”的水道网,这三种水道并不沟通,所留下的陆地的每一点的任意靠近处总有暖水、冷水及海水.换言之,这曲线全部地邻接着三个区域:暖水湖及其水道、冷水湖及其水道、海及其水道.

尽管上述图形是怎样奇特,在“平面区域的边界”的意义下这图形还应算

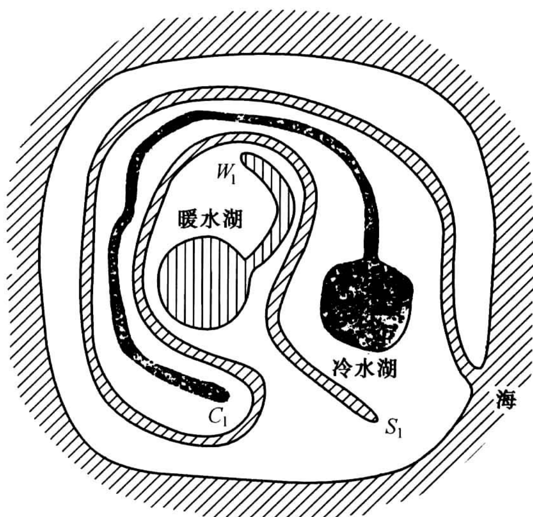


图 18.70

作曲线.

所考察的例指出，在这一般的观点下曲线的构成是相当复杂的. 第一个满意的广义的曲线(仅是平面曲线) 的定义是卓越的德国数学家康托①所给出，他是独特的数学科学 —— 集合论 —— 的创立者.

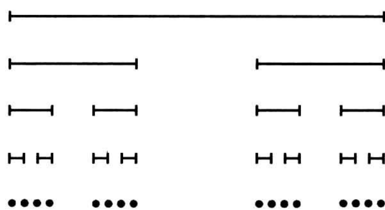


图 18.71

为了引向康托的曲线定义，先介绍一重要概念：每一个展布在平面上的集合如果它是闭的、连通的且有界的(展布在平面的有限部分上，不伸向无穷远)，则称这集合为平面连续统，并且认为连续统包含着多于一个的点(单点图形不作为连续统). 开区间(不闭)、抛物线(分支伸向无穷远)、两个同心圆周组成的图形(不连通) 都不是连续统. 圆周、圆盘、三角形、图形(b)(图 18.64) 和田曲线都是连续统. 显然，不是所有的连续统都是曲线，圆盘就不是曲线. 正是这样，展布在平面上的图形，如果它包含着内点，那就不是曲线. 我们称不包含内点的平面连续统(作为展布在平面上的集合) 为康托曲线. 也可以这样说，平面连续统 X 称为是康托曲线，如果它的任何点的充分靠近处总有不属于 X 的点. 这同样地意味着，在平面上去掉连续统 X 后所留下的集合(是一开集，因 X 是闭集) 以 X 为它的边界. 因而康托曲线是这样的连续统，它是平面上某些开集的边界. 在

① 可以注意图 18.63，18.64，18.68 及 18.72 的构造有共同之处，都用着所谓康托方法(连续地弃去一部分). 这些构造方法都起源于“康托完备集”. 这集合可由线段(0,1) 弃去其中间的三分之一，再弃去其余下各部分的，中间的三分之一，继续无限回而得到(图 18.71). 这集具有卓越的性质，常在构造及证明中应用它.

这意义下,用“曲面的边界”作为曲线的描述可以认为成功的.

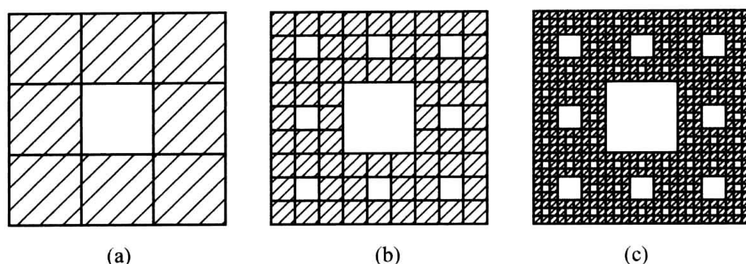


图 18.72

作为结尾,引进一康托曲线的例,这是波兰数学家希尔宾斯基(Sierpinski)给出的(图 18.72).把正方形分成九个相等的小正方形,去掉其中心的一个(图 18.72(a)).把余下的八个小正方形各分成九个相等的小正方形,各去掉其中心的一个(图 18.72(b));.把图 18.72(b) 中余下的 64 个小正方形作同样的处理(图 18.72(c)),无限地继续这一方法,在极限情况可得到曲线 C(通常称为希尔宾斯基铺盖).

图形 C 的卓越性质在于它是一“包罗万象的”曲线,有下列定理:

对于任何康托曲线 K,总可在图形 C 内找到这样的子集合与曲线 K 同胚.换言之,任何康托曲线,经过拉长、弯曲(不使断裂或粘叠)总能安放在图形 C 上.

但是,一般说,不是所有“曲线”都可以安放在图形 C 上的.前面的例(“房屋和井”)的曲线就不是“平的”(不能同胚于平面上的某一图形),也就是不能安放在平面上,当然更不能安放在图形 C 上.这里还指出,在三维空间存在着这样的图形 D,在它上面可以安放着任何“曲线”(在“曲线”一词的最一般理解下),而且图形 D 本身也就是这种广义曲线.图形 D 的构造有些和图形 C 的构造类似;用图 18.73 表示它的构造过程(按奥地利数学家芒格(Menger)的描述形式).什么是“曲线”一词的“最一般”的理解呢?我们现在转入阐明这一问题.

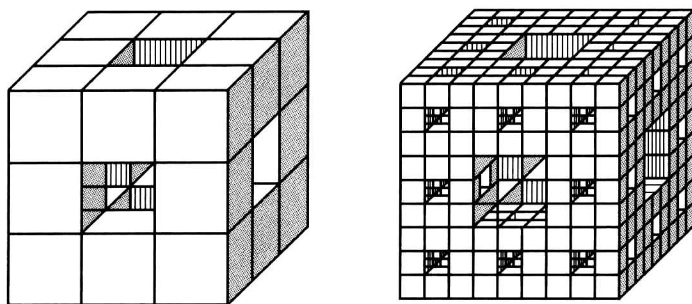


图 18.73

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

5.4 乌利松曲线

曲线的康托定义不是"内蕴的",它不容许按照集合本身的性质去确定这集合是否是曲线.这一定义仅对平面曲线适用,它不仅要利用我们所称为曲线的那集合的性质,并且还要用到这集合在平面上是怎样分布的.只要我们转向考察空间曲线,前面这许多研究都将失去意义.事实上,在空间,"区域的边界"是曲面,不是曲线[①].怎样去定义,什么是空间曲线呢?能否给出曲线的"内蕴的"定义,也就是这样的定义它同样好地适用于任何曲线而不依赖于这曲线是怎样分布(分布在平面或空间)的?

欧几里得还给出过曲线的第三个描述:"曲面有两个度量,曲线有一个度量,点一个度量也没有."从而,所希望给出的曲线的最一般的定义,可以说,曲线——是这样的图形(连续统)它有一个度量.这一描述并没有告诉我们什么内容,如果我们不知道怎样去严格地定义甚至怎样去描述度量的数目是什么?

许多数学家特别是庞加莱、布劳威尔、勒贝格企图定义什么是度量的数目或维数.但是这方面的最完整的结果属于卓越的苏维埃数学家乌利松(过早地死于1924年,只活了26岁).乌利松不仅以新的形式(不同于勒贝格-布劳威尔)定义了什么是度量的数目,并且建立了新的深刻的数学理论——维数论.特别是完全地解决了什么是曲线这一问题.维数论目前还在发展组成了拓扑学的重要分支.我们转入这一拓扑学分支的某些问题的概况.

6 维数

考察图18.74的图形,它是由圆盘、与之连接的线段及弧立着的点 P 组成.靠近点 M 所考察的图形是二维的(即有两个度量),靠近点 N 是一维的(有一个度量).点 P 和图形的其他部分是分开的,因而在这一点,图形是零维的(没有一个度量).所以,图形在它的不同点处具有不同的度量数目.

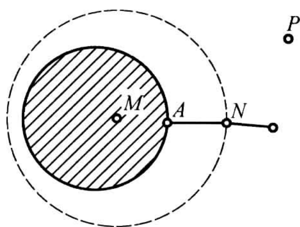


图18.74

6.1 乌利松的维数定义

为了定义图形在它的点处的度量数目(或维

---

① 可以试图定义空间曲线为:曲面是空间的区域的边界,而曲线是曲面上区域的边界(这里包含着上述欧几里得的"曲线"的描述).然而这样的描述是不能令人满意的.由这定义可以得出点和圆盘都可认为"曲线"(这里不再给出例子证实上述说法).

数),要引进隔离集概念.我们说,展布在图形 X 上的集合 A 隔离点 a 与点 b,如果在图形 X 上不存在这样的连通集,它包含着点 a 与点 b 而不与 A 相交.例如平面上的圆周隔离着这平面上在它的内部与外部的点.同样,空间的球面隔离着这空间的球面的内部与外部的点.对图 18.74 的图形 X,点 A 隔离着点 M 与点 N.在三维空间,点与点的隔离可以借助于二维图形来实现.(例如,借助于球面,图 18.75);而空间的一维图形(曲线)从外表上看已不可能隔离空间的点(因为可以"绕"过空间的曲线由一点走到另一点).在平面(二维图形)上,点连同和它靠近的点与平面上其余的点的隔离可以借助于曲线(例如圆周,图 18.76),也就是借助于一维图形来实现.在直线(一维图形)上,点连同和它靠近的点与直线上其他的点的隔离(图 18.77)可以借助于由两点 m,n 组成的图形,也就是借助于零维图形来实现.

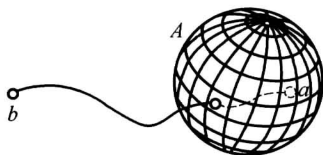


图 18.75

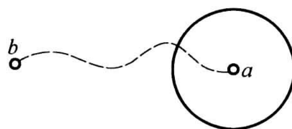


图 18.76

因而,可以说,在 $n(n=1,2,3)$ 维图形上,点的隔离可以借助于 $(n-1)$ 维图形(比原来的图形少一维)来实现.当然,隔离 n 维图形的点也可借助 n 维图形(例如可以用圆环隔离平面上的两点(图 18.78),但是小于 $(n-1)$ 维的图形是不能实现隔离的(例如不可能用一维图形——曲线——去隔离三维空间的两个点).

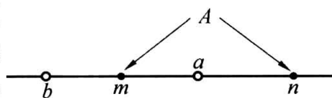


图 18.77

以上所述启发我们通过零维图形定义一维图形,通过一维图形定义二维图形,通过二维图形定义三维图形.①

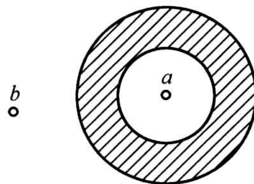


图 18.78

首先定义零维(没有一个度量的)图形.图形 X 称为是零维的,如果在这图形内不存在由多于一个点组成的连通集.粗略地说零维图形是"分散"的、孤立的,不连通的点.例如,一个点组成的图形是零维的,有限个点组成的图形是零维的.甚至存在着由无限个点组成的零维图形.

① 数学中考察的图形多于三维.甚至无限维,这种图形不仅在数学中而且在近代物理中起很大作用.在"四维欧氏空间"可以展开射影平面(或任何其他的单侧曲面)而使自身不相交.然而多于三维的空间不容许简单的直观表示,所以在本文中几乎不提到这种空间.

知道了零维图形可以定义一维图形.知道了一维,可以定义二维.一般说,知道了什么是$(n-1)$维图形我们可以定义什么是n维图形(下面是乌利松的定义).

设X是某一图形,a是它的点,图形X在点a具有维数n,如果满足下列条件:

(1) 在图形X内的点a的任意小邻域U内总能找到$(n-1)$维图形A,它把点a与图形X的处于邻域U外的所有点隔离.

(2) 对充分小的邻域U,这种隔离不可能借助于小于$(n-1)$维的图形来实现.

此外,图形X如果在它的某些点具有维数n,但在它的任何点不具有更大的维数,则称图形X是n维的.

下面举例说明上述(乌利松的)定义.直线是一维的:它不是零维图形(因为存在着由不止一点组成的连通集),而直线上点的隔离可借助于零维图形(图18.77).类似地,平面是二维的,空间是三维的.

图18.74的图形包含着点M,N,P,而且图形在这点处分别是二维、一维、零维的.所以就整个说,这图形是二维的.这例可以阐明前面的条件(2);如果取点M的比较大的邻域U(例如图18.74以虚线所画圆周的内部),则把点M与邻域外所有点隔离着的可以是零维图形(例如单点集N).但是不能说这图形在点M是一维的,因为对于M的充分小邻域,点M的隔离不能由零维图形来实现,必须是一维图形.

6.2　邦德里雅金图形

我们知道黏合圆环的内圆周上的对径点(处于直径端点的两点)则得到麦比乌斯带.如果把内圆周上分内圆周为三段相等弧的三点黏合成一点,所得图形称为模3的麦比乌斯带(图18.79).这图形可以描述如下:取一条卷了三圈的带(图18.80(a)),把它的下面的边黏合起来(图18.80(b)),即对下面的边说把三圈黏合为一圈.然后把这带的侧边PN,MN黏合(为了把点P,M黏合必须把侧边翻过去如图18.80(c)).这图形同胚于模3的麦比乌斯带.类似地,可以把圆环的内圆周上的四等分点,五等分点,…… 黏合起来,分别得到模4,模5,…… 的麦比乌斯带.任意模p(p是不小于2的自然数)的麦比乌斯带具有同胚于圆周的边界(就是所黏合的圆环的外圆周.图18.80(c)用粗线画出模3的麦比乌斯带的边界),从而模p的麦比乌斯带可以用来糊没某些曲面上切开的圆形洞.容易指出用模p的麦比乌斯带

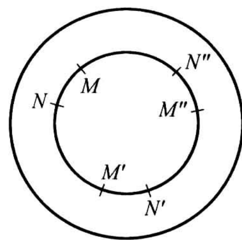
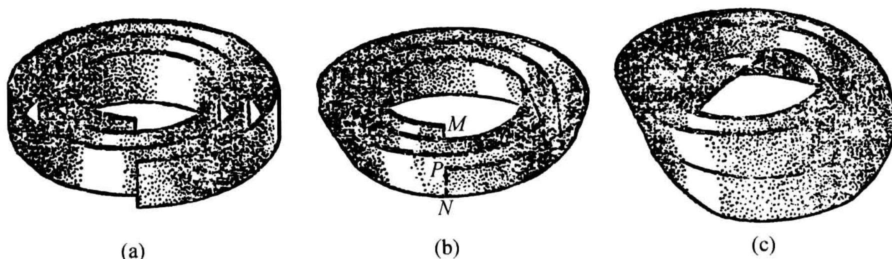


图18.79

糊圆形洞等价于把圆洞边缘的 p 等分点黏合起来.



(a) (b) (c)

图 18.80

取一正方形,切开一圆形洞,把这圆洞边缘上的 p 等分点黏合起来(也就是用模 p 的麦比乌斯带糊洞).然后再切开几个圆洞,同样地分别黏合相应的 p 等分点,接着再切开几个洞,再进行如上黏合(图 18.81).无限地进行下去,可以看到切开的洞、糊上的麦比乌斯带越来越小而且越来越密地布满在图形上,所得到的极限图形用 F_p 表示.这就是邦德里雅金图形.

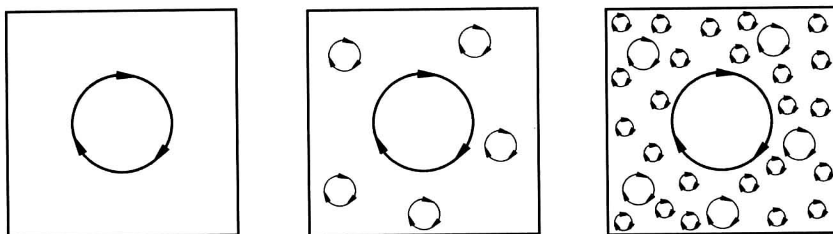


图 18.81

可以证明,对于任何的 $p = 2,3,4,5,\cdots$,图形 F_p 是二维的.现在变更上述构造过程,切开洞后,一次糊以模 p 的麦比乌斯带,另一次糊以模 q 的麦比乌斯带.这样交替地进行.模 p 的麦比乌斯带、模 q 的麦比乌斯带密布在图形上.这样得到的极限图形用 $F_{p,q}$ 表示,我们有:当 p,q 是互素时,图形 $F_{p,q}$ 是一维的.例如,我们都用模 3 的麦比乌斯带糊洞,则得到的图形 F_3 是二维的.如都用模 4 的麦比乌斯带糊洞,所得的图形 F_4 也是二维的.但是,如果我们交替地用模 3 的麦比乌斯带、模 4 的麦比乌斯带糊洞所得的图形 $F_{3,4}$ 却是一维的!

6.3 勒贝格－布劳威尔的维数定义

图形的维数也可以用另外方法定义,这一方法与杰出的法国数学家勒贝格及荷兰数学家布劳威尔的研究成果分不开.在严格地给出定义之前,先考察某些导引性的例子.

把直线分成段(线段),它们任何处都不是三段彼此邻接的(也就是没有点

属于三个线段):只要用等距离的点把直线分成许多线段(图 18.82),直线上的每一点或者属于一个线段,或者属于两个线段(当是线段的公共端点时).①

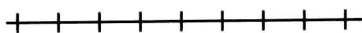


图 18.82

把平面分成小片使任何处都不是三片彼此邻接的,这是不可能的(虽然证明它不可能是很不简单的).当把平面分成正方形片时(用两组平行直线如图 18.83 所示),可以有四片彼此邻接(在正方形顶点处).在下述方法的意义下,这里的分法是“不经济”的.我们可以构造这样的分法,使任何处彼此邻接的片数不多于 3.这样的分法的例有“蜜蜂窠”(图 18.84)及“砌砖图案”(图 18.85).

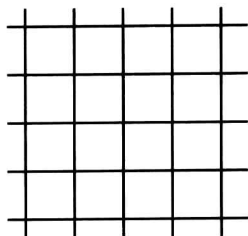


图 18.83

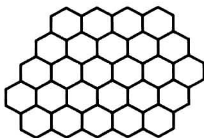


图 18.84

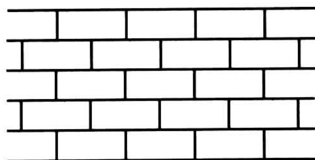


图 18.85

把三维空间分成小块使任何处彼此邻接的块数不多于 3,这是不可能的.然而,存在着这样的分法使任何处彼此邻接的块数不多于 4.为了得到这样的分法,例如可以砖平铺在地上如图 18.85 那样,然后铺上第二层.这时务使第一层砖的顶点不与第二层砖的顶点及棱碰到,如图 18.86 所示.图 18.87 表示邻接着的两层砖的分布(用黑线表示第一层,虚线表示第二层).从这个图形可看到任何邻接的块数不多于 4.

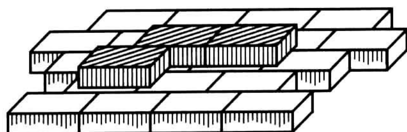


图 18.86

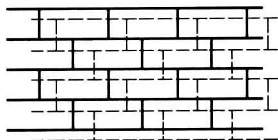


图 18.87

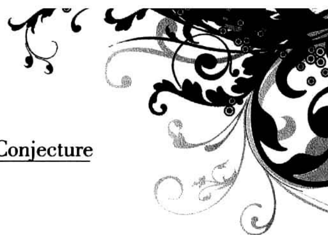
因而有:

一维图形(直线)可以分成段,使任何处邻接的段数不多于 2.

二维图形(平面)可以分成片,使任何处邻接的片数不多于 3.

① 这里和下面所考察的是分成闭的段或片,是包含着它们的边界的,所以这里的线段是把端点也算入的.





三维图形(空间)可以分成块,使任何处邻接的块数不多于4.

以上启示了给出维数新的定义的可能性.图形 X 称为是 n 维的,如果这图形可以由任意小的"片"组成(这种"片"是图形 X 的闭集),使任何处相邻接的"片"数不多于 $(n+1)$(也就是图形 X 中没有这样的点属于多于 $(n+1)$ 个"片"),但是,不能用数 n 去代替数 $n+1$.

可以证明,上述维数的定义容许有下述形式:图形是 n 维的,如果存在分法,把图形分成(任意小)的片,"涂"上 $(n+1)$ 种颜色,使任意两个相同颜色的片没有公共点.图18.88的平面涂上三种颜色(红、黄、青).

红	黄	青	红	黄	青
青	红	黄	青	红	
红	黄	青	红	黄	青
青	红	黄	青	红	
	黄	青	红	黄	

图18.88

乌利松的巨大功绩在于证明下述定理:上面考察的维数的两个定义是等价的.也就是说,如果图形按照第一个(乌利松的归纳的)定义具有维数 n,则按照第二个(勒贝格-布劳威尔)定义具有同样的维数 n.这两个定义并不是对所有拓扑空间都一致的,仅是对满足某些条件的空间(特别是度量空间)是等价的.对于这两个定义不等价的空间(非度量空间)的例曾为苏维埃数学家龙兹(Лунц)及洛库赞夫斯基(Локуциевскнй)所指出.

6.4 "邻居"

乌利松开创了拓扑学的大量的新篇章——维数论.维数是非常重要的拓扑不变量.这里叙述关于维数论的一些问题.首先,我们接触维数与"邻居数"的联系这一问题.最初发现考察这问题的是苏维埃数学家利赫金巴乌姆(Лихтенбаум).

把图形 X 分成许多"片",这许多片中与某一片 K 相邻接的片称为 K 的"邻居".对不同的片,它们的邻居的最大数称为所考察的分法(分成片)的毗邻数.如果图形 X 可以分成任意小的片使毗邻数是 S,而不可能分成任意小的片使毗邻数是 $S-1$,则称 S 是图形 X 的密度.

图18.84及图18.85所描述的分法的毗邻数是6,由于所分成的小片可以选取得任意小,所以平面的密度不多于6.但是,也许平面的密度实际上小于6,而是5呢?如果是这样的话,那么平面可以分成片使每一片的邻居的个数不多于5,平面的这样的分法是否可能?这问题的完全解决还是最近(1950年)洛库赞夫斯基及波尔金斯基(Болтянский)的工作.后者还证明,一般地,任何二维图形(满足某些当然的条件)具有密度不小于6.特别是平面的密度等于6,英国数学家斯通(Stone)在1953年证明,任何二维图形具有密度不大于7.所以,二维图

形的密度必须等于 6 或 7(例如具有三页的书的形状的图形的密度等于 7).

6.5 拓扑积概念

我们还要考察拓扑空间的"积" 的维数. 首先通过例研究"拓扑积" 这一新概念.

(1) 平面上任何点可由点对(一对点)(x, y) 给出, x 是横坐标轴上的点, y 是纵坐标轴上的点. 所以, 平面可以考察为所有可能的点对(x, y) 的集合, 这里 x 是图形 X(直线) 的点, 而 y 是另一图形 Y(也是直线) 的点.

(2) 考察如图 18.89 的圆环. X 是横截圆环的线段(它的延长线通过圆环中心), Y 是圆环的内圆周. 为了确定圆环上的任意点, 只要确定线段 X 的点 x 及圆周 Y 的点 y: 通过 x 作一同心圆, 通过 y 作一横截线段, 它们的交点就是所要确定的圆环上的点. 所以, 圆环可以考察为所有可能的点对(x, y) 的集合, 这里 x 是图形 X(线段) 的点, y 是图形 Y(圆周) 的点.

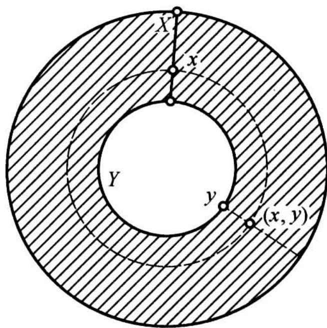


图 18.89

(3) 考察如图 18.90 的圆环面, X 是经圆(是通过圆环面的轴的半平面截圆环面所得的圆周), Y 是纬圆(垂直于轴的平面截圆环面所得的圆周). 为了确定这圆环面的点, 只要确定经圆 X 上的点 x 及纬圆 Y 上的点 y: 通过点 x 作一纬圆, 通过点 y 作一经圆, 它们的交点就是所需要确定的. 所以, 圆环面可以考察为所有可能的点对(x, y) 的集合, 这里 x 是图形 X(圆周) 的点, y 是图形 Y(另一圆周) 的点.

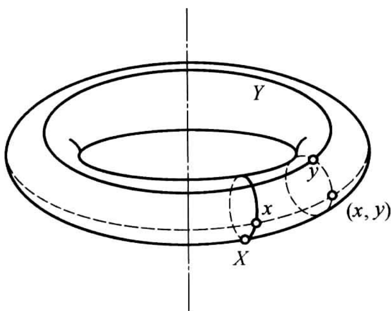


图 18.90

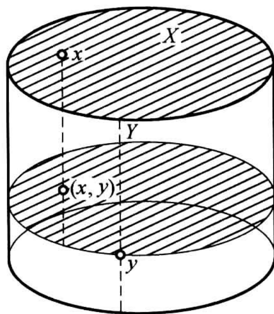


图 18.91

(4) 考察如图 18.91 的圆柱体. 用 X 表示它的上底, 用 Y 表示它的一条母

线.为了确定圆柱体上的点,只要确定圆盘 X 上的点 x 及线段 Y 上的点 y:通过点 x 引平行于母线 Y 的线段,过点 y 引平行于上底的圆盘,它们的交点就是所要确定的.所以,圆柱体可以考察为所有可能的点对 (x,y) 的集合,这里 x 是图形 X(圆盘)的点,y 是图形 Y(线段)的点.

在前面例中所得到的图形称为图形 X,Y 的拓扑积:平面是直线与直线的拓扑积,圆环是线段与圆周的拓扑积,圆环面是圆周与圆周的拓扑积,圆柱体是圆盘与线段的拓扑积.在一般情况,图形 Z 称为图形 X 及 Y 的拓扑积,如果 Z 可以考察为所有可能的点对 (x,y) 的集合,这里 x 是图形 X 的点,y 是图形 Y 的点.注意,上面只讲到怎样的点(点对)组成了图形 Z,但图形 Z 的作为拓扑空间没有涉及.下面的规定使图形 Z 成为拓扑空间:点 (x,y) 算作靠近图形 Z 的集合 M,如果对点 x 的(图形 X 中的)任何邻域 U 及点 y 的(图形 Y 中的)任何邻域 V,总能找到点 u 属于 U 及点 v 属于 V 使点 (u,v) 属于集合 M.

当然,图形 Z 的每一点对应着某点对 (x,y),而不同的点对对应着图形 Z 的不同的点.考察图 18.92 的球面,画有赤道 X 及零度的经线 Y.为了确定球面上的点,只要指出这一点的地图坐标(也就是赤道上的点 x 及零度经线上的点 y).为此,可以通过这一点作经线交赤道于 x,作纬线交零度经线于点 y.(x,y) 就是这一点的地球坐标.但是我们不能由此而得出结论:球面是赤道(圆周)与经线(同胚于线段的图形)的拓扑积.事实上,设 x,x' 是赤道上不同的两点,而 n 是经线的上端点(北极),则不同的点对 (x,n) 及 (x',n) 对应着球面上同一个点(北极).

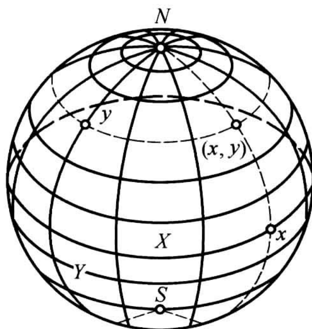


图 18.92

6.6 拓扑积的维数

现在转入拓扑积的维数问题所考察拓扑积的前面三个例子中,图形 X,Y 都是一维的,它们的拓扑积(平面、圆环、圆环面)都是二维的.如果一个图形是二维的,另一个图形是一维的(圆盘与线段),则它们的拓扑积是三维的(圆柱体).从这些例及另外一些简单的例(平面与直线的拓扑积是三维空间,平面与平面的拓扑积是四维空间)似乎可以断言:拓扑积的维数等于它的"相乘"空间的维数的和.虽然这一断言没有被证明,长时期来,人们不去怀疑它的正确性.邦德里雅金于 1930 年发表的文章引起了科学界的轰动.他指出两个二维图形的拓扑积可以是三维的(如按前面的断言应是四维的).当然,并不是任意两个二维图形的拓扑积是三维图形,但是邦德里雅金首先指出可以找到这样两个二

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

维图形它们的拓扑积是三维图形.

邦德里雅金所采取的图形就是前面6.2所描述过的他自己构造的图形 F_p，如果 p,q 是互素的数，则邦德里雅金图形 F_p 与 F_q 的拓扑积是三维图形. 这一定理的证明并不简单.

注意邦德里雅金得到这一结果是把两个不同的(不是同胚的)图形 F_p 与 F_q"相乘". 如果 p,q 不是互素的(具有公因子)，甚至如果 $p=q$，则图形 F_p,F_q 的拓扑积已不是三维的，而是四维的(符合前面维数相加的断言). 自然产生如下的问题：在邦德里雅金的例中两个图形的不同是否本质的？能否存在着这样的(二维)图形"乘以"自己本身后得到的拓扑积是三维图形？这一问题的答案是波尔金斯基在1949年给出的，他构造了新的图形 $P_m(m=2,3,4,5,\cdots)$. 这些图形 P_m 是二维的且具有如下性质：对任意的 m,n(特别是 $m=n$)图形 P_m,P_n 的拓扑积是三维图形.

7　基本群

这里我们研究一个重要的拓扑概念，它紧密地联系着群这一代数概念. 我们指出，对任一图形，更明确些，对每一拓扑空间可以构造确定的群 —— 这空间的基本群. 尽管这一概念表面上很抽象，它有一系列具体应用和实践特征(例如纽结的理论).

在给出基本群的定义之前，回到前面关于曲线概念所考察的道路概念.

7.1　道路及其形变，同伦道路

设点 x 在某连通的①拓扑空间中连续移动，描出以 x_0 为起点 x_1 为终点的道路 h. 这道路连续地形变，保持端点 x_0 及 x_1 不动，在空间 X 中移动(图18.93). 道路 h 形变为另一道路 h' 的过渡位置在图 18.93 中用细线画出.

注意，在形变过程中，道路可以自身相交，可以多次地通过空间 X 的同一点. 例如在图 18.94 中描出的形变，道路 h 转变为另一道路 h'，这最后结果并不是自身相交的.

设道路 h 及 h' 能借助于形变把一个转变为另一个，则称这两个道路是互为同伦的，表示为 $h\sim h'$. 再一次强调指出，我们所考察的道路的形变仅是保持端点不变的情况.

① 这里及以后称图形(或空间)是连通的，是指对它的任意两点可用道路连接起来(这性质也称"弧连通性"，具有弧连通性的图形也具有连通性，其逆不真).



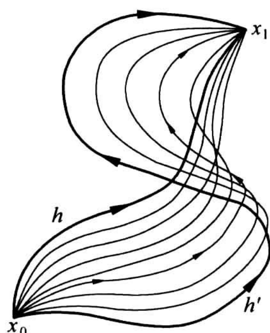


图 18.93

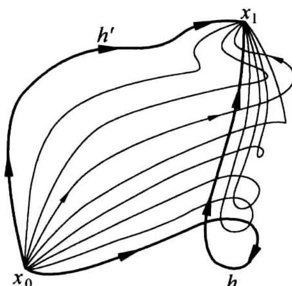


图 18.94

容易知道,在圆盘上任意两个具有公共端点的道路 h_1 与 h_2 是彼此同伦的.事实上,我们可以把道路设想为可伸缩的橡皮线,它蜿蜒地展布在圆盘的内部.如果我们固定端点后容许它自由变动,就紧缩为联结端点 x_0, x_1 的直线段.所以在圆盘内的联结点 x_0 及 x_1 的任意道路都同伦于直线段 x_0x_1,从而圆内的任意两条道路彼此同伦.

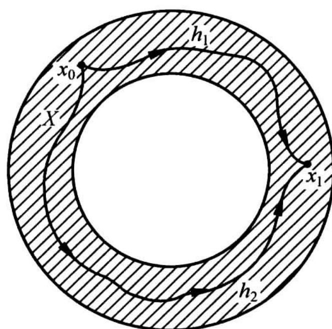


图 18.95

在圆环内部,每两个具有公共端点的道路并不都是彼此同伦的.例如图 18.95 所画的道路 h_1 及 h_2,由于圆环的"洞"起着阻碍作用,不可能在环的内部由一个道路形变为另一道路.

7.2 道路的积,道路的同伦类

设一条道路 k 的起点恰好是另一道路 h 的终点,则所得到的先经由 h 再经由 k 的道路称为道路 h 与 k 的积,记作 hk,按此定义,只有这样的道路可以定义它们的积.就是第一道路的终点恰好是第二道路的起点.

设 $h \sim h', k \sim k'$,并且 h 的终点与 k 的起点重合,则 $hk \sim h'k'$.

在空间 X 中固定某一点 O,我们考察起点和终点都是 O 的道路,对这种道路来说任意两个道路总是可以相乘的.①

把彼此同伦的道路归入同一类,从而把所有的道路分类.我们考察道路类,不考察道路本身.这道路类称为同伦类.在圆盘内,所有道路都是同伦的,所以

① 注意,一般说道路 hk 与 kh 是不同的,甚至通常彼此不同伦.

只存在唯一的同伦类.在圆环内(图 18.96),道路 h,k 并不是同伦的.道路 k 并不环绕着圆环的"洞",可以收缩为一点,而道路 h 环绕着这洞,不可能收缩为一点.所以,h,k 分别属于不同的同伦类.

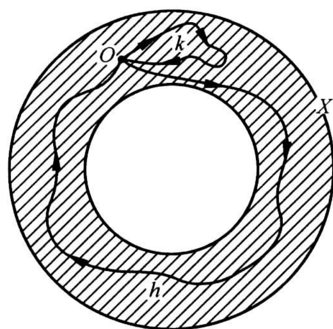


图 18.96

因而,在空间 X,选定了固定点 O,考察以点 O(称为基点)为起点及终点的闭路的不同的同伦类.所有这些同伦类所成的集合用 $F(X)$ 表示.这些类可以相乘:设 h 属于第一个类,k 属于第二个类,把 h,k 相乘,称乘积 hk 所属的类为这两个类的乘积.如果,我们在两个类中分别取 h',k' 代替 h,k,如上所指出的,所得乘积 $h'k'$ 与 hk 彼此同伦,也就是属同一类.所以两个类的乘积的定义为两个类所确定,与所选取的"代表"无关.

7.3 基本群

在上节末,已在集合 $F(X)$ 内,也就是在空间 X 的道路(以口为起讫点)的同伦类所成的集合内定义了乘法.我们指出关于这乘法运算,集合 $F(X)$ 形成群.下面简单地叙述这群是怎样形成的.

设 k 是任一道路,h 是可以收缩为一点的道路(图 18.97(a) 及(b)),两条道路都是以 O 为起讫点.不难指出 $hk \sim k$(图 18.98),$kh \sim k$.所以,所有可以收缩为一点的道路所成的类可以用 e 表示.我们得到对任意类 a,$ae = ea = a$.采用群论的术语,类 e 称为单位元①.此外,设 a 是某一类,h 是这类的某一道路,则把与 h 方向相反的道路记作 h^{-1}(图 18.99).不难验证,道路 hh^{-1} 及 $h^{-1}h$ 收缩为一点.用 a^{-1} 表示 h^{-1} 所属的类,则有 $aa^{-1} = a^{-1}a = e$.这就是说对 $F(X)$ 中的每一元 a 都存在逆元.最后,对于任意三道路 h,k,l,我们有 $(hk)l \equiv h(kl)$.所以在 $F(X)$ 中关于乘法是可结合的.所以集合 $F(X)$ 关于所引进的乘法运算是一个群,这群称为空间 X 的基本群.

注意,群 $F(X)$,一般说是不可变换的,即 $ab \neq ba$.

① 在下面,群 $F(X)$ 的单位元将不用符号 e 表示,而用更习惯的符号 1 表示.

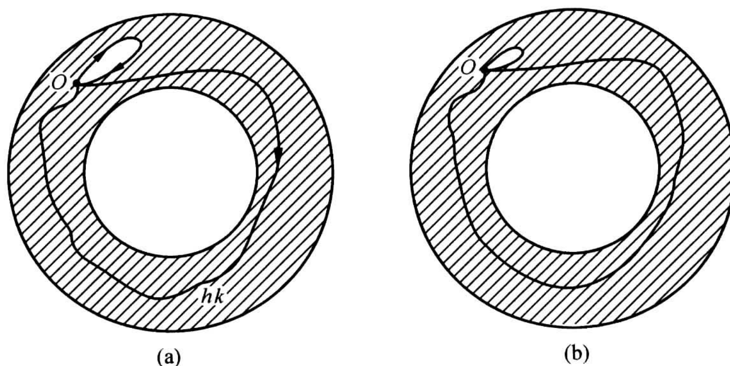


图 18.97

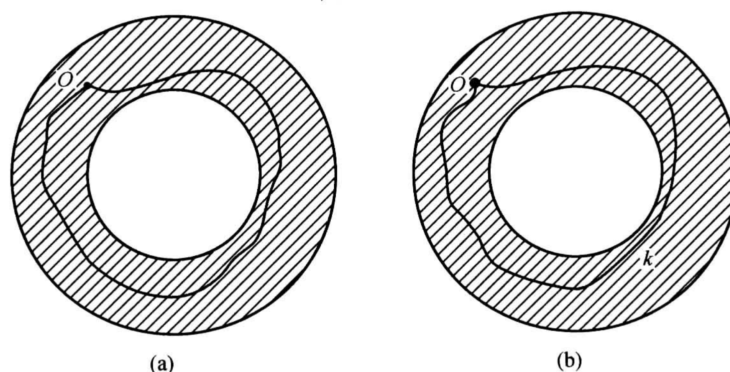


图 18.98

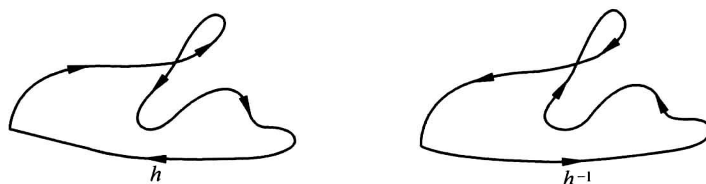


图 18.99

7.4 基点的变换

上面引进基本群定义时依赖于任意选取的基点 O.然而不难证明,由不同基点 O 及 $O*$ 所构造的空间 X 的基本群是彼此同构的.事实上,设 w 是连接点 O 与 $O*$ 的任意道路(图 18.100).对于每一个以 O 为起讫点的任意道路 h.使 $O*$ 为起讫点的道路 $h^* = w^{-1}hw$(w^{-1} 是与 w 方向相反的道路)与之对应.不难验证,这是关于点 O 的道路类与关于点 $O*$ 的道路类间的一一对应.此外,这一对应是同构的:由关系式 $h^* = w^{-1}hw, k^* = w^{-1}kw$ 可知道路 $h^*k^* =$

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

$w^{-1}hww^{-1}kw$ 与道路 $w^{-1}hkw$ 同伦. 这就是由 h^*k^* 确定的类对应着由 hk 确定的类. 这指出:借助于不同的基点构成的基本群是彼此同构的.

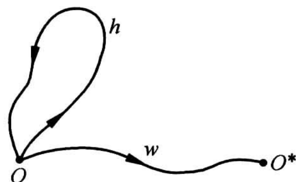


图 18.100

可以更一般地指出:设连通空间 X 与 Y 是同胚的,则它们的基本群 $F(X)$,$F(Y)$ 是同构的.这样,作为代数结构的基本群是空间 X 的拓扑不变量.

如果两个空间的基本群不是同构的,则这两个空间不是同胚的.

7.5 例

稍后一些,我们将对充分广泛的一类图形(通常称为多面体)指出找寻基本群的方法.现在考察一些直接找出基本群的例子.

(1) 圆盘的基本群(如我们所知)是平凡的,也就是仅由一个单位元组成的.

为了便于理解下面的例可以注意下述事实:在光滑曲面上,任何道路(甚至类似于充满整个平面的皮亚诺曲线那样的道路)都可以形变为不覆盖整个平面的光滑道路.

(2) 二维球面(三维空间的球的表面)的基本群也是平凡的.事实上,球面上的任意道路可以形变为不充满整个球面的光滑道路.对所考察的道路,可在球面上去掉一点,使这一点不为所考察的道路所经过,则所得图形同胚于圆盘,而在圆盘内道路可以收缩为一点.

具有平凡的基本群的空间称为单连通空间.圆盘、球面都是单连通空间.此外,直线、平面,甚至 n 维球面都是单连通空间.回忆基本群的定义,我们可以直接定义单连通性:"如果这空间内的任意闭的道路可以收缩为一点,这空间称为单连通的".

(3) 圆周的基本群不是平凡的,它同构于所有整数按加法构成的群(这样的群称为无限循环群①).事实上,按某方向(算作正的)环绕圆周一圈的道路记作 a,相反方向的道路记作 a^{-1},则 a^n 表示环绕圆周 n 圈的道路,按正的方向,n 取正整数,按相反方向,n 取负整数. a^0 表示着可以收缩为一点的道路.

为了更清楚地考察圆周上的不同道路,我们描绘相应的点的运动图象.按道路的定义,点的每一位置对应着参数 t 在单位线段($0 \leqslant t \leqslant 1$)上的值(例如 t 表示时间).另外,每一点的位置对应着圆周上的确定的坐标角 ϕ(为方便起

① 这样的群也称为自由群,因为它不包含任何关系式以确定元素间的关系.



见,起点 O 的坐标角算做 0).用横坐标轴表示 t 值,用纵坐标轴表示 ϕ 值,可作出 $\phi = \phi(t)$ 的图象($\phi(0) = 0$).

如果点均匀地环绕圆周 n 圈,我们得到道路 a^n;它的图象可以用直线段表示(图 18.101).但是点在圆周上可以取复杂的形式运动,例如在运动时可以几次地改变方向.图 18.102(a) 所示的道路的图象见图 18.102(b).圆周上的任意闭路的图象总是连接点(0,0) 与$(1,2n\pi)$的,因为通过这道路而回到点 O 时,坐标角是 2π 的倍数.

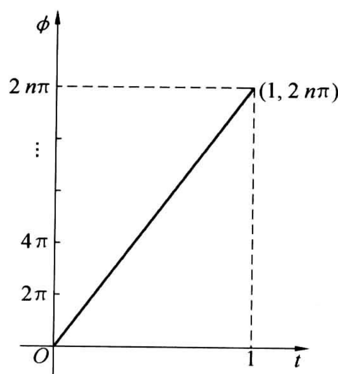
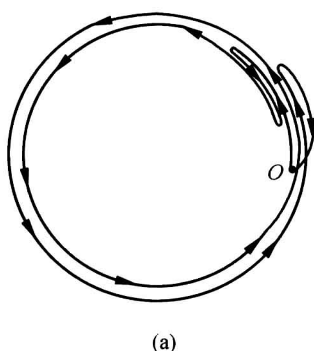
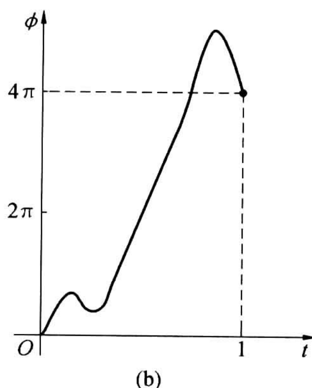


图 18.101

如果某些道路的图象联结点(0,0)与$(1,2n\pi)$,则称 n 为所考察道路绕圆周的回转次数.不难看到,任何道路 f 它完



(a)



(b)

图 18.102

成n 圈必同伦于道路 a^n, 为此,可画出道路 f 及道路 a^n 的图象,让 f 的图象上的点垂直地投影到 a^n 的图象上(图 18.103).f 的图象就形变为直线段,所以道路 f 与道路 a^n 同伦.环绕不同次数的道路彼此不同伦.

所以,圆周的基本群的元素与整数集合一一对应.余下的,注意当道路相乘,则它们的回转次数相加,所以圆周的基本群同构于整数加群.

(4) 射影平面的基本群.取任意点 O 作为起点及某一通过点 O 的直线,取定方向,用 a 表示.我们考察直线 a 作为闭路,这闭路由 O 出发按射线的方向(前取定的方向)引向无穷远点,然后通过直线的另一部分的射线回到点 O.在这平面上把直线 a 环绕点 O 旋转 180°,这是道路 a 到道路a^{-1} 的形变.用 α 表示由道路 a 确定的基本群的元素,我们有 $\alpha = \alpha^{-1}$,也就是 $\alpha^2 = 1$(引用基本群的

单位元 e，则有 $\alpha^2 = e$）. 不难指出，在点 O 的任意闭路可以在射影平面上“投影”到直线 a，也就是同伦于道路 a^n. 但是由于道路 a^2 可以收缩为一点（因为 $\alpha^2 = e$），所以任何道路或者同伦于道路 a，或者可以收缩为一点. 可以指出道路 a 不能收缩为一点. 从而得到，射影平面的基本群是由二个元素组成：e 和 α（这里 $\alpha^2 = e$），也就是二阶的循环群.

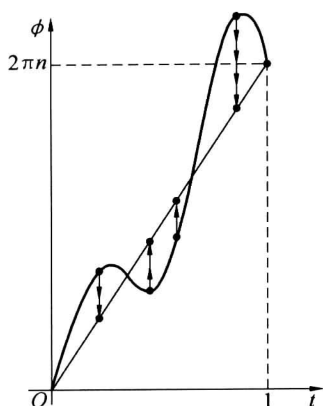


图 18.103

7.6 胞腔剖分与多面体

任意拓扑空间的基本群的计算是如此复杂的问题，在这里未必能给出任何具体的计算方法. 我们限制在比较简单的图形——多面体，它在拓扑学（特别组合拓扑学）中起着重要作用，按其本质说是研究的基础对象. 我们首先指出什么图形叫做多面体.

前面我们考察过网络，它是描绘在曲面上的. 网络的棱把曲面剖分成许多部分——面，每一部分同胚于圆盘. 最简单的例子是凸多面体借助于它的棱所得的剖分. 在图 18.104(a)，(b)，(c) 中，面与面间的相对分布是比较复杂的.

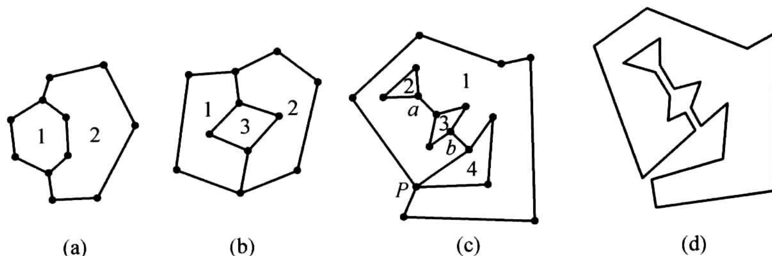


图 18.104

在图 18.104(c) 中，标以数字 1 的面自身相交于棱 a，b 及顶点 P. 如果按照它所有的棱切开，则这面变为同胚于圆盘的片（图 18.104(d)）.

我们考察更一般的图形（不仅是曲面），它可以剖分为多角形. 也就是这样的图形，在它上面可以“画”上网络，把图形剖分成分离的片，这些片同胚于圆盘. 这种片称为面，而网络的元素称为棱与顶点. 面与面连接时有非常复杂的情况（例如图 18.104(a)，(b)，(c)）；并且容许某些棱连接三个，四个，…… 面（不必像有边的曲面那样连接一或两个面），也容许这样的棱，它不与任何面连接（如图 18.105），沿着这种棱切开图形是不可能的.

能够借助于网络剖分为同胚于圆盘的片的图形称为二维多面体，而把分裂

所得的棱、顶点、面的集合称为这个多面体的胞腔剖分(或简称剖分).自然,同一多面体存在着不同的剖分.还可以考察一维多面体(不包含面);一维多面体的每一个剖分就是网络,因为它是由顶点和棱组成.零维多面体是由个别的点组成.也有三维,四维,……,n维多面体,三维多面体的剖分是由顶点、棱、面及体组成.如果按三维多面体的所有的顶点、棱、面切开,则剖分为分离的三维片——体,它同胚于球.

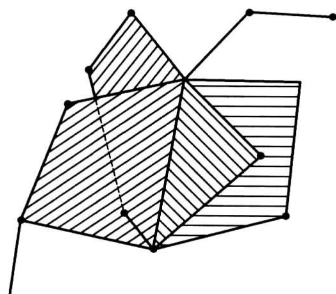
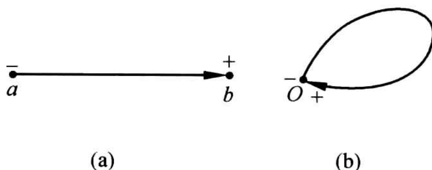


图 18.105

剖分的每一个元素(即顶点,棱,面,……)也称为胞腔(零维,一维,二维,…… 胞腔).多面体是最简单的图形.如我们所知,在拓扑学中也研究比较复杂的图形,但不是多面体.

7.7　面的同伦边界

胞腔剖分的某些棱称为是定向的,如果在它上面可以给定方向.代替箭头,可以用“+”“-” 表示方向;终点用“+”,起点用“-”.如果这棱是闭的(端点重合),则在棱的运动开始部分标以“-”,在运动终止部分标以“+”.如图 18.106所示.



(a)　　(b)

图 18.106

我们还要讲到某些面的边界的线路方向.在凸多边形的情况,面的边界线路方向是明显的;在较复杂情况(如图 18.104(c)),规定如下:我们知道沿着所有的棱切开后,每个面转变为同胚于圆盘的片.按某一方向环绕这片的边界一周.然后把切开处“黏合”,则方才环绕着片的线路给出了按胞腔边界的运动方向,这就是面的边界线路.当以相反方向环绕片的边界,则面的边界线路也变为相反方向.可以给出线路定义的另一描述:面的边界线路不一定就是这面的边界,而是充分接近边界且无处相交的线路(如图 18.107(a),(b)).

现在引进面的同伦边界的概念,这一概念在多面体的基本群的计算中起重要作用.

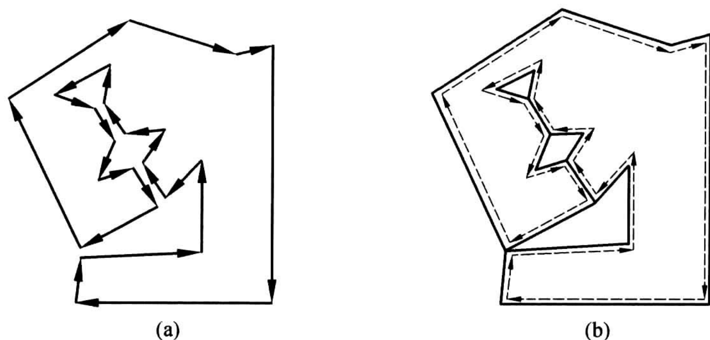


图 18.107

考察胞腔剖分的某些面,连接着这个面的所有的棱作为已经确定了方向并用文字($a,b,c,\cdots$)表示.我们用下述方法完成面的边界线路,并用代数式(单项式)表示它.如果线路从棱 a 开始我们首先写下 a 或 a^{-1} 以表示按箭头方向或相反方向通过这棱.如果下面要通过的棱是标以"d"的,则在 a 或 a^{-1} 的右边写上 d 或 d^{-1}(图 18.108),以此类推,完成整个线路,写下了某单项式.这单项式称为所考察面的同伦边界.

注意,如果所有的棱都已定向并用文字表示,则面的同伦边界的写出依赖于环绕的方向和由那一棱开始.对每一个面,我们可以任取一种同伦边界的记录(单项式).在所得的单项式内不能随便地改换因数的次序.

图 18.108 给出了胞腔 Ⅰ、Ⅱ、Ⅲ、Ⅳ 及所有棱的方向和标号.我们得到按逆时针方向的这些胞腔的同伦边界为:

Ⅰ:$adbc$,Ⅱ:$kh^{-1}g^{-1}fd^{-1}g$,Ⅲ:hl,Ⅳ:$l^{-1}k^{-1}$.

一般说,在同伦边界内,同一棱可以出现多于一次(如上述 g 在 Ⅱ 中的情况).

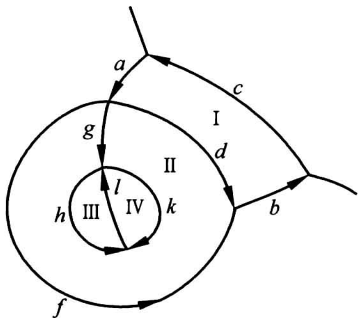
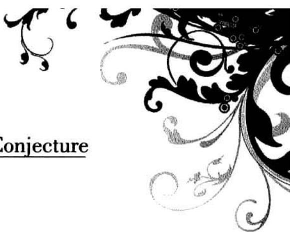


图 18.108



7.8　树,网络的最大树

现在引进下述概念.某些连通的网络(图 18.109) 称为树,如果从它的一个顶点到它的另外的任何顶点只能经过一条折线(由网络的棱组成).或者说,如果所考察的网络不包含任何闭折线.

在每一个连通的网络里可以选出(一般说,有不同方法) 最大的树.也就是这样的树,当加上所考察网络的任一棱后就不再是树了.关于构造最大的树,可以取网络的任一顶点,并由此出发逐步地添加(当可能的话) 网络的所有的棱,务使每次添加都得到树.(注意最大的树必须包含所考察图形的所有顶点)

取如图 18.110 所描的例.从顶点 a 出发按顶点 b 的方向进行,经过道路 abc 到达分支的点 c,这里有两个方向进行,向 d 或者向 e.向 d 的方向到 d 结束,因为添加线段 da 后构成闭路 $abcda$,不再成为树了.向 e 的方向可以到达道路 cef,在那里可按两个方向中的一个方向(例如向 g 的方向) 进行,在点 h 有两个方向,其一到 l 结束,另一到点 o.在点 o(这里由点 n 引来) 有三个方向:op, os, ot.最后得到如图 18.104 的树,它不包含线段 da, fl, jp, mk.

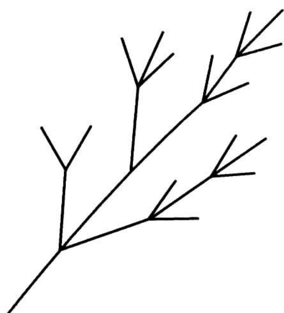


图 18.109

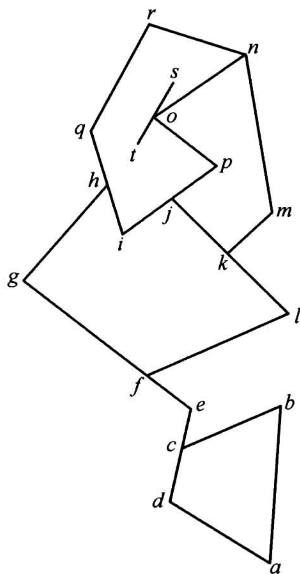


图 18.110

7.9　多面体基本群的计算方法

这里给出计算连通的多面体的基本群的简单方法,不加以证明.

设 P 是任何连通的多面体.取这多面体的任意的胞腔剖分,把这剖分的所

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

有顶点和棱组成的网络记作 G.在网络 G 中任意选取最大的树,而把构成这树的所有的棱都标以数字1.对网络 G 中其余的棱(不属于最大树的)我们设法给以定向并标以不同文字 $a,b,c,\cdots$.此外,对于这胞腔剖分的每一个面写出它的同伦边界(单项式),略去其中(单项式中)的等于1的因数(即标以数字1的棱).最后,取不属于网络 G 的最大树的棱 $a,b,c,\cdots$ 作为群的生成元,令所有的同伦边界(单项式)等于1,把得到的这些等式确定元素 $a,b,c,\cdots$ 间的关系(称这些等式为确定关系式或简称为关系式).这样得到的群同构于连通多面体 P 的基本群.这种计算基本群的方法是很方便的.

7.10 例

1.圆周.可以取由一顶点与一棱组成的胞腔部分(图18.111).最大树由一个顶点组成,所以这棱 a 是基本群的唯一的生成元,而且没有确定关系式(因为没有二维胞腔 —— 面),所以圆周的基本群是自由循环群.

2.双纽线(图18.112)的基本群是两个生成元 a,b 的自由群(没有确定关系式).

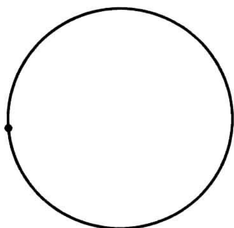


图18.111

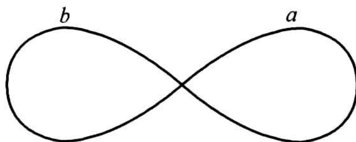


图18.112

3.球面 S(二维).可以取由一个顶点,一个面组成的胞腔剖分.刺破了一点的球面同胚于圆盘(考察为没有边界的).最大树仅由一点组成.由于没有棱,所以它的基本群没有生成元,从而是平凡的.

4.射影平面.如图18.113所示的剖分是由一个圆盘和它的边界组成,这边界分为两个半圆 r_1,r_2 具有一致的方向.黏合圆周上的对径点得到射影平面.这时棱 r_1 与 r_2 黏合为一个棱 r.这样,射影平面的胞腔剖分是由一个顶点、一个棱 r 和一个面组成.这个面的同伦边界是 $r\cdot r=r^2$,所以射影平面的基本群是由一个生成元 r 及关系式 $r^2=1$ 构成,也就是二阶循环群(也称模2的剩余类群).

5.模 m 的射影平面.设有 m 个点把圆周分为 m 等份,如果把圆周上具有这性质的 m 个点黏合为一点(图18.114),则得到模 m 的射影平面①.圆周分为 m

① 换言之,模 m 的射影平面可以由用模 m 的麦比乌斯带糊球面上切开的洞而得到.

段相等的弧 $r_1, r_2, r_3, \cdots, r_m$，且具有一致的方向．当黏合时这 m 条弧合为一弧 r．我们得到胞腔剖分由一个顶点、一个棱 r 及一个面组成，并且这面的同伦边界是 $r \cdot r \cdot \cdots \cdot r = r^m$，故得到仅有的确定关系式 $r^m = 1$，所以得到的基本群是 m 阶循环群（即模 m 的剩余类群）．

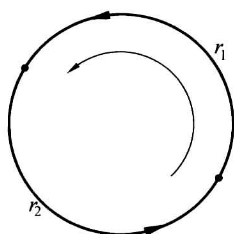


图 18.113

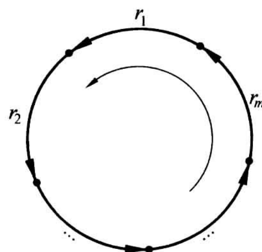


图 18.114

6．环面．在环面上画经圆和纬圆，它们相交于点 o，得到环面的胞腔剖分由一个顶点 o，两个棱 a（纬圆），b（经圆）及一个面组成（图 18.115），沿经圆、纬圆切开，则环面变为矩形（图 18.116）．这里唯一的面的同伦边界为 $aba^{-1}b^{-1}$，所以环面的基本群是由两个生成元 a, b 及关系式 $aba^{-1}b^{-1} = 1$ 形成，即 $ab = ba$，也就是说这基本群是具有两个生成元的自由阿贝尔群．

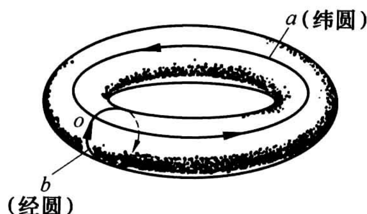


图 18.115

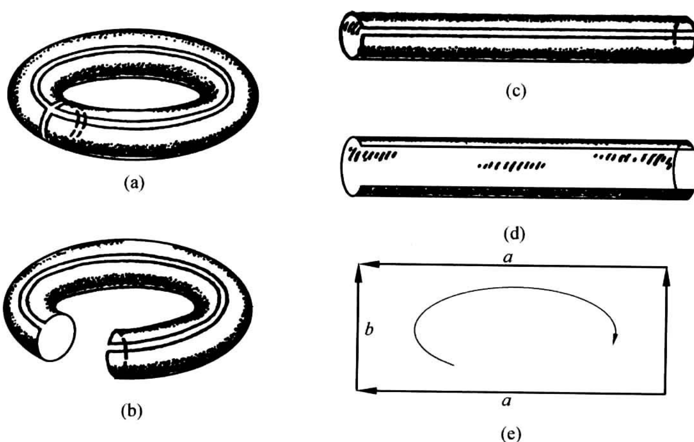


图 18.116

7．如果在环面上挖去一洞如图 18.117 所示，沿经圆、纬圆切开时不再得出矩形，得到的是五边形（图 18.118）．黏合图 18.118 标以相同文字的线段，则变五

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

边形为有洞的环面,也就是柄.棱 c 是这柄的边界.取同样的两个柄,沿着它们的边界黏合起来则得到双环面.换言之,把图 18.119 标以相同文字的线段黏合起来则得到双环面,也就是定向曲面 P_2.



图 18.117

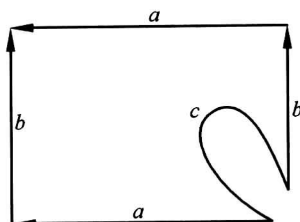


图 18.118

挖了洞的双环面可以黏合图 18.120 的多角形得到.再粘上一个柄(图 18.121)得到具有三个柄的球面,也就是定向曲面 P_3.一般说,定向曲面 P_m(也就是具有 m 个柄的球面)可以由 $4m$ 角形得到,即黏合图 18.122 中标以相同文字的边而得到,从而曲面 P_m 的胞腔剖分可以表示为:一个顶点,$2m$ 个棱 $a_1, b_1, a_2, b_2, \cdots, a_m, b_m$, 及一个面. 这面的同伦边界为 $a_1b_1a_1^{-1}b_1^{-1}a_2b_2a_2^{-1}b_2^{-1}\cdots a_mb_ma_m^{-1}b_m^{-1}$(图 18.122).这样,曲面 P_m 的基本群是由 $2m$ 个生成元:$a_1, b_1, a_2, b_2, \cdots, a_m, b_m$ 及关系式

$$a_1b_1a_1^{-1}b_1^{-1}a_2b_2a_2^{-1}b_2^{-1}\cdots a_mb_ma_m^{-1}b_m^{-1} = 1$$

确定.注意当 $m \geqslant 2$ 时这个群不是阿贝尔群(例如 $a_1b_1 \neq b_1a_1$).

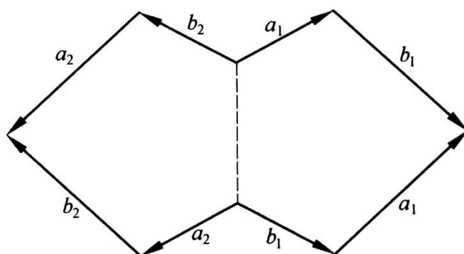


图 18.119

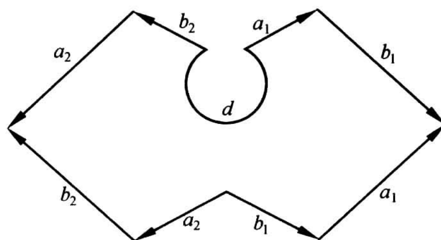


图 18.120

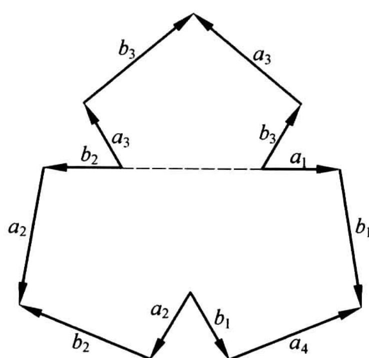


图 18.121

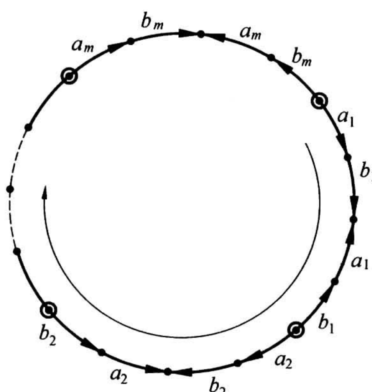


图 18.122

7.11 纽结和纽结群

各种各样的纽结具有许多应用.仅登山者为各种不同目的而采用的纽结已达几十种拓扑类型.纽结也为海员、纺织工人、针织品工人等所采用.

最简单的纽结如图 18.123 所示,我们称它为简单纽结①.

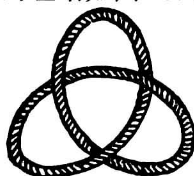


图 18.123



图 18.124

简单纽结有两种类型:左的和右的.借助于镜照映射可以把左纽结变为右纽结.关于右纽结的命名可作如下理解,给纽结的线以确定的一致的方向(图 18.125).纽结的部分(b)按右旋方向构成环孔,而部分(a)以右手螺旋的前进方向进入环孔.

有趣的是人们常习惯于或者总打右纽结,或者总打左纽结.

下面是较复杂的二重结(图 18.126)②,我们不能把它和所谓海员结(图 18.127)相混淆,因为它们的组成不同.

---

① 图 18.124,打结的那条"绳"的两端并不接合,从而这结总是可以"解开"的,所以在拓扑学中考察的纽结或者是在以无穷远点为端点的线上,或者是在封闭的线上.我们将考察的是后一种情况.

② 图 18.126(a) 给出通常二重结的图案(在封闭的线上);图 126(b) 给出这结的实际形式.图 18.127 的情况相类似.

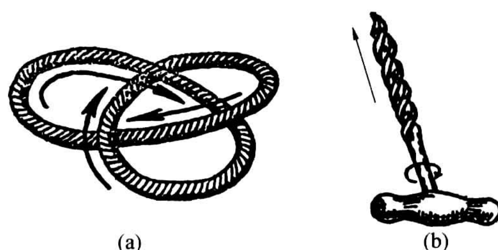


图 18.125

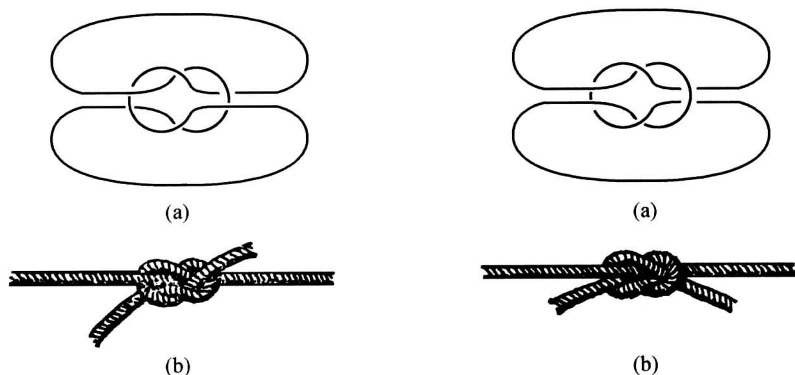


图 18.126　　　　图 18.127

三维空间拓扑学的重要问题是下述所谓纽结问题.

三维空间的简单闭曲线 L_1 和 L_2(例如两闭折线) 在什么情况下,可以认为在空间是拓扑一致分布的?

图形的"一致分布" 不够确切,下面精确地定义两图形合痕的重要概念.拓扑空间 R 的两个集合 L_1 和 L_2 称为在 R 中是拓扑一致分布的,或在 R 中是合痕的,如果存在着 R 到自身的拓扑映象把集合 L_1 映成集合 L_2.

这样,纽结的问题在于找出空间 E_3 中的两闭曲线 L_1 与 L_2 合痕的条件.实际上,例如打结的闭曲线与不打结的闭曲线(图 18.128) 并不是拓扑一致分布的(也就是不是合痕的).这在直观上是明显的,但是怎样证明呢?

纽结问题至今没有完全解决.然而知道了不同纽结的拓扑不变量使能在一系列情况下区分纽结.最简单的纽结不变量是纽结群,是以余空间 $E_3 \setminus L$ 的基本群 $F(E_3 \setminus L)$ 定义的(三维空间去掉属于纽结 L 的点外的集合称为余空间 $E_3 \setminus L$).纽结 L 的纽结群记作 $G(L)$, $G(L) = F(E_3 \setminus L)$.如果纽结 L_1, L_2 的纽结群不同构,则纽结 L_1 与 L_2 不是合痕的(这可由基本群是拓扑不变量推得).

这里给出纽结群的实际计算方法.把闭折线 L 投影到水平的面上,在平面上得到的折线(折线 L 的投影) 可能是自身相交的,然而我们可以假设至多具有二重交点(必要的话,可稍稍移动折线中的线段),没有其他特异(可能出现的



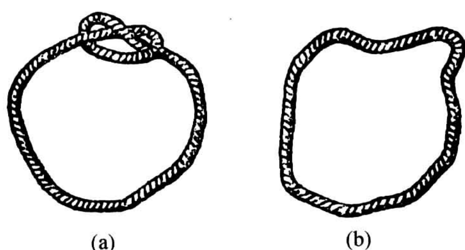


图 18.128

特异如图 18.129 的上半部分所示,这图的下半部分排除了这些特异).所以我们可以认为投影上的每一个交点恰好是折线中两个线段的交点.我们让交点处的两线段中的在下的线段在交点处中断,在上的线段仍保持接通,并且用光滑(无角)的曲线代替这折线(图 18.130).这样的纽结图形称为纽结的正规投影.

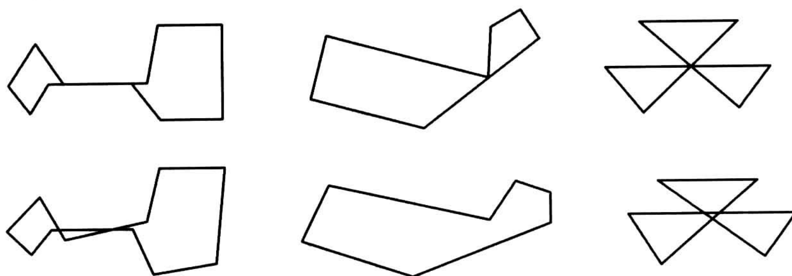


图 18.129

现借助正规投影计算纽结群.把投影分为 n 个中断的弧,表示为 a_1, $a_2,\cdots,a_n$,而曲线 L 上对应的部分记作 $\alpha_1,\alpha_2,\cdots,\alpha_n$.此外,选定曲线 L 的方向并在弧 $a_1,a_2,\cdots,a_n$ 上标以箭头(图 18.131).为了构造基本群,在空间取一点 o,使位于曲线 L 外边,并从这一点引道路包围弧 αk,以右手螺旋的方向环绕着弧 αk(图 18.132).这样的道路用 xk 表示.道路 $xk(k=1,2,\cdots,n)$(正好是这些道路的同伦类)是基本群 $F(E_3\setminus L)$ 的生成元,也就是纽结群的生成元.不难理解,余空间的任意道路同伦于这 n 个道路的组合.余下的要建立这些生成元间的关系式.

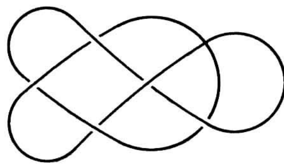


图 18.130

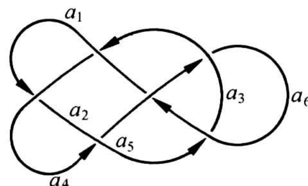


图 18.131

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

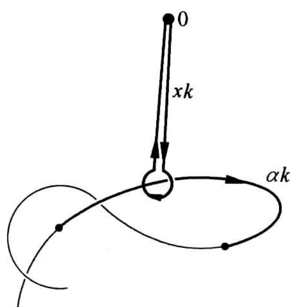


图 18.132

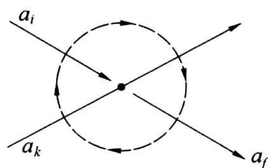


图 18.133

为此,考察投影上的任意一个二重点.令集中在这一点的弧如图 18.133 所示.环绕这二重点作一顺时针方向的不大的圆周.用下面的方法写下某些单项式:如果某些弧进入这圆周,则相应的道路 x 的指数取 $+1$,如果某些弧从这圆周内引出,则相应的道路 x 的指数取 -1.我们沿着圆周环绕这二重点,依次地遇到集中在这点的四条弧(如图 18.133 的射线),并且自左到右依次写下四个因数(x 的幂)的乘积.让这乘积等于 1.例如图 18.133 的二重点,可得关系式

$$x_i xk^{-1} x_j^{-1} xk = 1$$

(不难想象,道路 $x_i xk^{-1} x_j^{-1} xk$ 在空间 $E_3 \setminus L$ 中同伦于零).对每一个二重点,写下这样的关系式,所有这些关系式确定了生成元 $x_1, x_2, \cdots, x_n$ 间的关系.

7.12 例

1.考察纽结,它的投影如图 18.134 所示.这纽结的群具有两个生成元 x_1,x_2(对应着投影上的弧 a_1, a_2)具有两个关系式(在二重点 P_1, P_2)

$$x_1^{-1} x_1 x_1 x_2^{-1} = 1, x_2 x_1^{-2} x_2^{-1} x_2 = 1$$

这两式都归结为 $x_1 = x_2$.所以这纽结群是自由循环群,也就是和圆周的群一样.如图 18.134 显然所考察的纽结与圆周是合痕的.

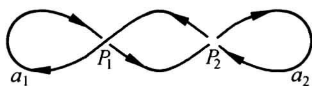


图 18.134

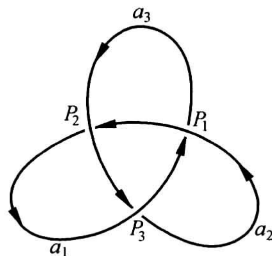


图 18.135

2.简单纽结(图 18.123)的投影如图 18.135 所示.在二重点 P_1, P_2, P_3 处的

关系式(关于生成元 x_1, x_2, x_3)有

$$x_2x_1x_3^{-1}x_1^{-1} = 1$$

$$x_3x_2x_1^{-1}x_2^{-1} = 1$$

$$x_1x_3x_2^{-1}x_3^{-1} = 1①$$

设 $x_1 = x, x_2 = y$,由上述第二式得 $x_3 = yxy^{-1}$.代入第一式(或第三式)得 $xyx = yxy$.所以这群是由两个生成元及唯一关系式 $xyx = yxy$ 形成.

可以指出,这群不是可交换的,从而不与无限循环群同构,所以图 18.123 的纽结与圆周不是合痕的(特别是不能不切断而解开这纽结).

3.上述推理可以不加改变地适用于比较复杂的纽结,这种纽结是由某些简单闭曲线自身相互交错而成.我们限制于考察三个圆相互交织如图 18.136 所示(圆周间得以如图分布自然允许他们在空间稍微弯曲).不难看到,如果拉破了其中一个圆,另外两个就分开了.然而我们不能拆开任何一个圆使它们彼此分离,所以所描出的纽结与分离地分布的三个圆不是合痕的.要证明这事实,我们只要证明这纽结群不能同构于由三个生成元构成的自由群(这是分离地分布的三圆的群).图 18.136 的纽结群有六个生成元:$a_1, a_2, \cdots, a_6$,联系着六个关系式(对应着交点 $P_1, P_2, \cdots, P_6$),这些关系式是

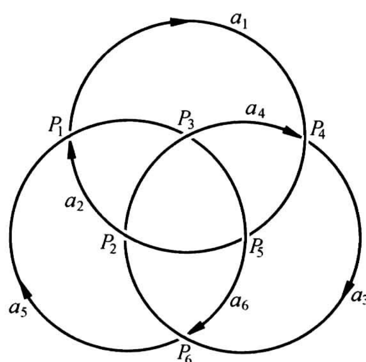


图 18.136

$$a_1^{-1}a_5^{-1}a_2a_5 = 1, a_4^{-1}a_2a_3a_2^{-1} = 1$$

$$a_4^{-1}a_6^{-1}a_4a_5 = 1, a_1a_3^{-1}a_1^{-1}a_4 = 1$$

$$a_1a_6^{-1}a_2^{-1}a_6 = 1, a_6a_3a_5^{-1}a_3^{-1} = 1$$

设 $a_1 = x, a_3 = y, a_5 = z$,代入第一、四、六式中得

$$a_2 = zxz^{-1}, a_4 = xyx^{-1}, a_6 = yzy^{-1}$$

代入其他三式中得关系式

$$xyx^{-1}zx = yzy^{-1}xy = zxz^{-1}yz$$

所以图 18.136 的纽结群具有三个生成元 x, y, z,联系着上述关系式.余下的,我们要借助于这样的事实:"具有 n 个生成元联系着至少一个不平凡的关系式的群不能同构于具有 n 个生成元的自由群"这一定理的证明(并不简单)见关于群

① 原译者注,三个关系式与图 18.135 不符.按该图所示相应的关系式是 $x_2x_1x_2^{-1}x_3^{-1} = 1, x_3x_2x_3^{-1}x_1^{-1} = 1, x_1x_3x_1^{-1}x_2^{-1} = 1$.经简化后,结果相同.

论的书.

8 同调群

这里我们研究重要的拓扑不变量——同调群,按其本质来说早已为庞加莱所引进.同调概念的应用不仅在拓扑学本身而且在数学的其他分支:代数、分析、几何等.为了了解这末尾的内容,要求读者具有群论的最简单性质的知识.第一部分带有引言的性质,不要求预备知识.

8.1 引言——同调群的直观描述

我们称某些曲面 X 上的环路是指这曲面上的有向闭曲线或某些这种曲线的总和.例如,图 18.137 环面上的定向经圆或定向纬圆,或它们的总和都是环路.

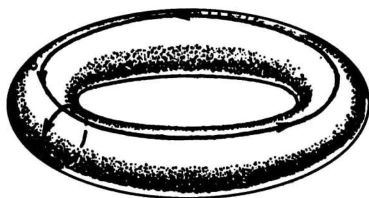


图 18.137

曲面上的环路可以分为两类,有些是界住着某些区域的(也就是作为某些区域的边界),还有一些不能界住任何区域的.这一概念,初看起来简单、明显(例如,平面上的圆周或其他自身不相交的闭曲线界住着它内部的区域).事实上要求明确化时,正是会出现不明显性.

我们作为曲面 X 是双侧的,选取曲面的某一片 U,在它的一侧涂以任何颜色.对于 U 的边界——曲线 z 选取这样的方向,使得面对着 U 的涂颜色的一侧沿着曲线 z 按这方向移动时左边看到的点属于 U(右边看到的点不属于 U).在这情况,我们把这定向曲线 z 作为小片 U 的边界(或者说 z 界住着 U).显然,如果我们在 U 的另一侧涂以颜色,则它的边界是同一曲线但取相反的方向,这一环路记作 $-z$.

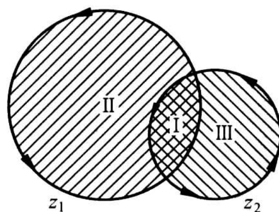


图 18.138

界住区域的环路并不像初看时那么简单.在图 18.138 所画同一平面上的两个环路 z_1,z_2.每一环路界住着它内部画有细线条的区域.这两个环路的和(表示为z_1+z_2)界住着这样的区域,这区域是由取小片 Ⅰ 两次,小片 Ⅱ、Ⅲ 各一次组成(环路 z_1+z_2 所界住的区域 $U\equiv 2\cdot$ Ⅰ $+1\cdot$ Ⅱ $+1\cdot$ Ⅲ),所以环路所界住的区域有时必须取曲面 X 的某些小片而带有确定的系数.在这引言中我们不停留在这一问题上,在后面研究同调群时可

以明确地阐述.

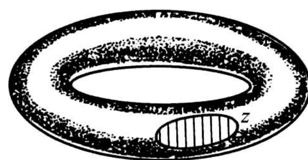


图 18.139

图 18.139 所示环面上的环路 z 界住着画有细线条的区域,而环面上的经圆或纬圆或把两者合起来所成的环路都不能界住着任何区域.

显然,界住着区域的环路可以在任何曲面上引进.在曲面上的每一"充分小"的圆周都是界住着区域的环路.不能界住任何区域的环路遇到较少.例如,在环面上存在着不界住区域的环路,而在球面上就不存在.从而我们可以得出结论,球面不同胚于环面.图 18.140 所示双环面上的环路界住着画有细线条的区域.这种为环路所界住的区域称为胞膜,它是紧绷在这环路上的.

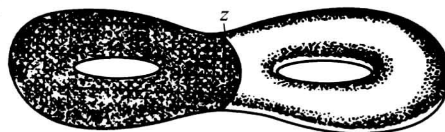


图 18.140

我们称这种界住着区域的环路为"非本质的",而把不界住任何区域的环路列入"本质的"环路.我们可以认为两个不同的环路 z_1,z_2 是"非本质地彼此不同",如果环路 z_1 与环路 $-z_2$ 的和形成非本质的环路(在这情况,也称环路 z_1,z_2 彼此同调).例如图 18.141(a) 所示环面上的两个经圆具有相同的方向,图 18.141(b) 上环路 z_1 与环路 $-z_2$ 的和(环路)界住着画有细线条的胞膜.所以图 18.141(a) 的环路 z_1,z_2 彼此同调.图 18.142 所示的曲面上的环路 z_1 与 z_2 也彼此同调,因为这里由 z_1 与 $-z_2$ 的和(环路)界住着画有细线条的区域(胞膜).

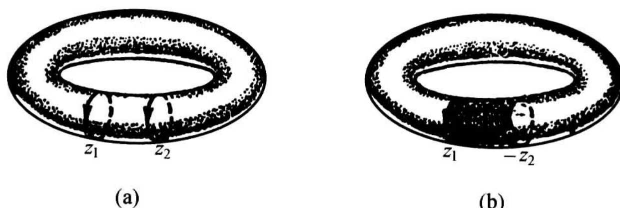


图 18.141

今后我们感兴趣的问题是在给定的曲面上可以引进多少彼此不同的本质的环路?显然,在环面上"环路的存储"多于球面上,在双环面上"环路的存储"又多于环面上(图 18.143 的双环面上的四条环路:p_1,p_2,m_1,m_2,其中两条是经圆,两条是纬圆,它们都是本质地不同的,也就是彼此不同调).这里某些不加以定义的术语"给定曲面所有环路的存储",要在同调群建立以后才精确定义.

给定曲面所具有的环路的集合可以分为不同的环路类(称为同调类),把所有彼此同调的环路归入一类.例如,在球面上只有一个同调类,所以它的所有环

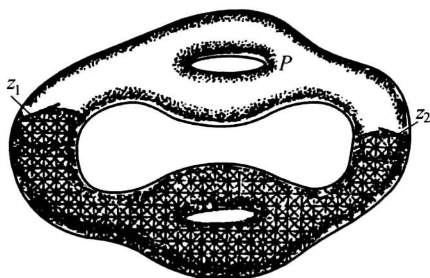


图 18.142

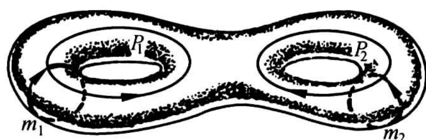


图 18.143

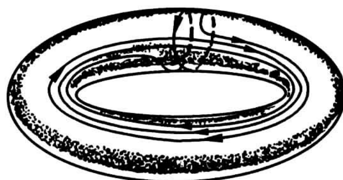


图 18.144

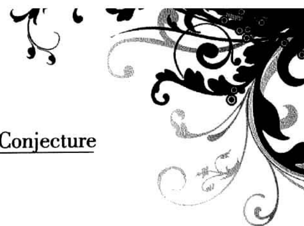
路都彼此同调的(它们都界住着某些区域).在环面上存在着无限多个同调类,但是它们都可以用两个基本同调类——经圆类和纬圆类表示.例如按纬的方向环绕环面三次按经的方向环绕环面二次的所有环路(图 18.144)形成的同调类可以表示为 $2m+3p$,这里 m 是经圆,p 是纬圆.一般说,环面上的任一同调类具有形式 $\alpha m+\beta p$,这里 α,β 是整数,所以环面上的所有同调类形成群(可以像二元一次多项式那样进行加减),并且这个群具有两个生成元.在双环面上(图18.143)有四个生成元:m_1,m_2,p_1,p_2,每一同调类具有形式 $\alpha_1m_1+\alpha_2m_2+\beta_1p_1+\beta_2p_2$,所以双环面的"同调群"比环面的"同调群"具有较"多"的元素,从而这两个曲面不同胚.

这些例阐明某些曲面的所有同调类的集合自身是一群,称为同调群.例如,球面上的同调群是平凡的,环面上的同调群是具有两个生成元的自由阿贝尔(Abel)群,双环面上的同调群是具有四个生成元的自由阿贝尔群.

怎样在一般情况构造同调群呢?在所考察曲面上所有的一般环路组成阿贝尔群(环路可以进行加、减).所有界住着区域的环路也组成群——环路群的子群(因为两个界住着区域的环路的和与差仍是这样的环路).我们希望"收集"彼此同调的环路于一类;特别应"收集"所有界住着区域的环路("同调于零"①)于一类.

① 用代数的术语,我们应按界住着区域的环路子群取环路群的商群.事实上.两个环路 z_1 与 z_2 当且仅当属于商群的一个元素时,它们的差 z_1-z_2 属于界住着区域的环路子群,也就是它们彼此同调.换言之,商群的元素就是同调类.商群本身称为所考察曲面的同调群.





我们前面所考察的例仅是在二维图形(曲面)上"一维"的环路(曲线).在拓扑学中要考察不同维数(零维、一维、二维等)的环路,它们所界住着的区域可以是多于二维的.还注意为某些非本质的环路所界住着的"区域"(就是紧绷在这环路上的胞膜)可能展布在所考察的图形的"较厚"处.例如,设 X 是三维空间为环面所界住着的部分(也就是 X 是环面上及环面内的点所成集合,称为环体),则环面上的经圆成为图形 X(环体)的非本质的环路(图 18.145),因为在 X 中这环路上可以"紧绷着胞膜"(也就是以这环路为边界的小片),这正是图 18.145 环体的截面.相反地,这图形上的纬圆仍是本质的环路,它不能界住着 X 的任何胞膜.

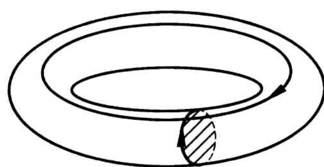


图 18.145

两个环路的同调概念与它们间的形变概念有很重要的联系:设有两个环路 z_1, z_2,如果其中一个能通过形变移动到另一个上去(这样两个环路同伦),则两个环路彼此同调.这可以在图 18.141(a) 看到.但是相反叙述不成立,图 18.142 所示曲面 P 上的同调环路不是同伦的,把 P 上的环路 z_1 移动到环路 z_2 时为曲面 P 上界于两个环路间的"洞"所阻碍.所以对两个环路的同调性说,同伦性是它的充分条件但不是必要条件.

8.2 定向,关联系数

一条棱有两个方向,规定方向的棱称为定向棱.在拓扑学中不仅对一维胞腔(棱),对二维,三维,…… 胞腔也给以定向.至于点(零维胞腔),不给以定向.

二维胞腔的定向是由它的边界(曲线)的环绕方向给出.由于二维胞腔的边界有两种不同的环绕方向(顺时针方向和逆时针方向),所以每一个二维胞腔有两种不同的定向.所谓给定了多边形的定向意味着对它的边选定了这样的方向从多边形的内部观察这些边(棱)时,它们的方向是一致的(图 18.146).

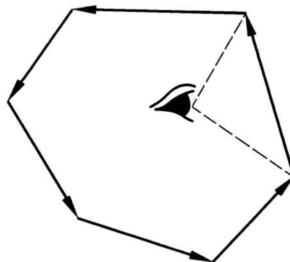


图 18.146

同样可以给出三维胞腔的定向,为了对多面体定向,可以给它的所有的面以这样的定向,使从多面体的内部观察这些面时,它们的定向是一致的(都是逆时针方向,如图 18.147;或都是顺时针方向).如果从多面体的内部观察它的所有的面时,这些面都是按顺时针方向定向的,这时视线的方向恰和右手螺旋(图 18.148)的前进方向一致.多面体的这

种定向有时称为按右手螺旋定向.相反情况称为按左手螺旋定向.

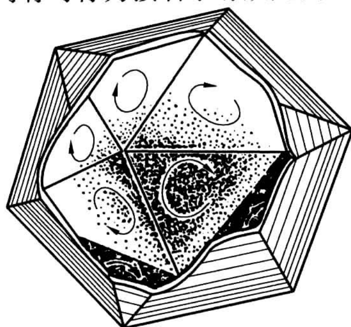


图 18.147

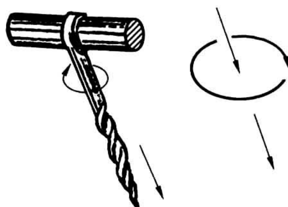


图 18.148

现在介绍二维胞腔的关联系数这一重要概念.首先规定定向棱对它的顶点的关联系数.每一棱有两个顶点:起点和终点.如果这棱不是闭的(即它的起点和终点不重合),则这棱对它的起点的关联系数算做 -1,对它的终点的关联系数算做 $+1$.在闭棱情况(起点与终点重合)系数 $+1$ 与 -1 对销,这时对这顶点的关联系数算做 0.

现在定义定向面对定向棱的关联系数,在沿着定向面的边界环行时所通过的棱的方向与环行方向一致,则定向面对这定向棱的关联系数为 $+1$;如果所通过的棱的方向与环行方向相反,则相应的关联系数为 -1.如果在环行中某棱往复通过几次,则关联系数为每一次通过的关联系数的代数和.从上述定义即可得出:假使组成了面 g 的同伦边界(按定向面 g 的环行方向),则这个面对某些定向棱 r 的关联系数等于 r 包含在这边界里的幂的指数的和.

以麦比乌斯带为例.我们知道它可以表示为圆环形式而黏合它的内圆周上的对径点.如图 18.149 取线段 pq 连接外圆周上的点 p 及内圆周上的点 q,而 q' 是 q 的对径点.点 q,q' 把内圆周分为相互黏合的两部分(如图箭头所示).这两部分都用 a 表示,定向棱 pq 用 b 表示,外圆周用 c 表示.这时麦比乌斯带的胞腔剖分包含着两个零维胞腔 p 与 q,三个一维胞腔:a,b,c,一个二维胞腔 τ.胞腔 τ 的同伦边界(由点 p 出发对外圆周作逆时针方向)为

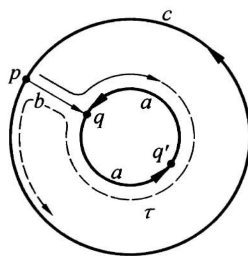


图 18.149

$$ba^{-1}a^{-1}b^{-1}c = ba^{-2}b^{-1}c$$

所以胞腔 τ 对 a,b,c 的关联系数是 $-2,0,+1$.



射影平面可以表示为胞腔剖分形式包含着一个顶点 a,一条棱 r 和一个面 g;面对棱的关联系数为 ± 2(依赖于面和棱的定向).模 m 的射影平面可以由一点、一棱、一面表示为胞腔剖分形式;面对棱的关联系数为 $\pm m$.环面可以由一点、二棱 r_1,r_2 及一个面 g 表示为胞腔剖分形式:面对每一棱的关联系数为0.定向曲面 P_m(m 个柄的球)可以由一点、$2m$ 条棱及一面表示为胞腔剖分形式;面对每一棱的关联系数为 0.此外,在前面四个例中,棱对顶点的关联系数都是0(因为所有的棱都是闭的).

8.3 链及其边缘

不闭的定向棱 ab(记作 r)对起点 a 的关联系数为 -1,终点 b 的关联系数为 $+1$.我们定义这棱 r 的边缘为

$$\Delta r = b - a$$

这里 Δ 是"边缘符号"(Δr 读作"胞腔 r 的边缘"),而位于右端的两个点的代数"和"有下列意义:在顶点前面的符号给出了对应的关联系数.当 r 是闭的胞腔时,我们有 $\Delta r = 0$(更详尽些,我们可以写作 $\Delta r = 0a$).

类似地可以定义二维胞腔的边缘:如果 g 是二维定向胞腔,则定义它的边缘 Δg 为包含在胞腔边界内的棱的"和"

$$\Delta g = \sum_i \varepsilon_i r_i \tag{①}$$

棱 r_i 所取的系数 ε_i 是面 g 对棱 r_i 的关联系数.例如在麦比乌斯带情况,如图18.149所描绘的,胞腔 τ 的边缘为

$$\Delta\tau = -2a + 0b + 1c = c - 2a$$

如果我们同意胞腔 g 取相反的定向时记作 $-g$,则我们有

$$\Delta(-g) = \sum_i(-\varepsilon_i r_i) = -\sum_i \varepsilon_i r_i = -\Delta g$$

因为当胞腔变换方向时,所有的关联系数都变号.

式 ① 的右边部分是带有某些整系数的定向棱的和.这个和,当然仅是形式的,因为我们没有可能引进两个棱的"加法".

每一个带有某些整系数的胞腔的(形式的)和称为链.链有零维的、一维的、二维的等随着我们所加的是怎样的胞腔而定.例如等式 ① 的右端是一维链.链可以相加、相减、乘以数.链的相加和多项式一样,当出现同类项时按通常的法则化简(回忆当变换胞腔 g 的定向时,所得的胞腔用 $-g$ 表示,从而对应着同一胞腔的两个不同定向的两项是同类项).

借助于加法运算,所有二维链(同样,一维链或零维链)的集合形成阿贝尔群;这群的零元素是这样的链,它的系数都是零.这个群称为二维(一维或零维)链群,记作 $L_2(P)$($L_1(P)$ 或 $L_0(P)$).这里 P 表示所考察的链所依据的那个剖

分,附标 2,1,0 表示所考察链的维数.

上面我们定义了胞腔的边缘,下面我们可以定义任意链的边缘. 设 $x_2 = \sum_i \alpha_i g_i$ 是二维链(α_i 是系数,g_i 是二维胞腔),则它的边缘 Δx_2 定义为

$$\Delta(\sum_i \alpha_i g_i) = \sum_i \alpha_i(\Delta g_i)$$

由于个别胞腔的边缘我们已能计算,上式的右边部分不难计算. 类似地可以定义一维(或三维)链的边缘. 任意零维链的边缘规定为 0.

不难相信两个(或几个)链的代数和的边缘等于边缘的代数和. 例如,设 x'_2 及 x''_2 是两个二维链,则

$$\Delta(x'_2 \pm x''_2) = \Delta x'_2 \pm \Delta x''_2 \qquad ②$$

事实上,设 $x'_2 = \sum_i \alpha_i g_i, x''_2 = \sum_i \beta_i g_i$(对所考察剖分的所有二维胞腔求和,系数 α_i, β_i 可以为零),则有

$$\Delta(x'_2 + x''_2) = \Delta(\sum_i \alpha_i g_i + \sum_i \beta_i g_i) =$$

$$\Delta \sum_i (\alpha_i + \beta_i) g_i = \sum_i (\alpha_i + \beta_i) \Delta g_i$$

$$\Delta x'_2 + \Delta x''_2 = \Delta(\sum_i \alpha_i g_i) + \Delta(\sum_i \beta_i g_i) =$$

$$\sum_i \alpha_i \Delta g_i + \sum_i \beta_i \Delta g_i = \sum_i (\alpha_i + \beta_i) \Delta g_i$$

边缘算子 Δ(把链 x_2 变为 Δx_2)是二维链群 $L_2(P)$ 到一维链群 $L_1(P)$ 的映象:如 $x_2 \in L_2(P)$,则 $\Delta x_2 \in L_1(P)$;并有和 $x'_2 + x''_2$ 对应着和 $\Delta x'_2 + \Delta x''_2$,所以这一映象是群 $L_2(P)$ 到群 $L_1(P)$ 的同态映象.(在一般情况,Δ 是群 $L_n(p)$ 到群 $L_{n-1}(P)$ 的同态映象)

8.4 边缘的基本性质

链的边缘运算具有如下性质:对于任意链 x 有

$$\Delta(\Delta x) = 0 \qquad ③$$

(边缘的边缘等于零). 对于一维链 x,这一关系是显然的,因为 Δx 是零维链,它的边缘(按定义)为零. 下面就二维链阐明关系式 ③.

首先考察二维胞腔 g,它是一多边形,它的边缘没有任何“黏合”情况. 按照胞腔 g 的定向环绕 g 的边界,并在这多边形的边上用箭头表示环绕的方向(图 18.150). 这时,组成边缘的棱都有确定的定向,并且所有关联系数都等于 1. 从而 $\Delta g = r_1 + r_2 + \cdots + r_n$,这里 $r_1, r_2, \cdots, r_n$ 是多边形 g 的边,它们的定向如图 18.150 所示. 因为在多边形 g 的每一个顶点有一条棱进入,也有一条棱由此引出,所以在链 $\Delta(\Delta g) = \Delta r_1 + \Delta r_2 + \cdots + \Delta r_n$ 中每一顶点出现两次:一次取系数 $+1$(作为进入的棱的端点),一次取系数 -1(作为引出的棱的端点),从而



$\Delta(\Delta g)=0$.

类似地,在比较复杂的多边形它的边缘产生黏合的情况(图 18.151). 在环绕胞腔的边界时,注意在 g 的每一顶点进入棱的数等于引出的棱的数,即加数与减数相等,所以出现在 $\Delta(\Delta g)$ 中的每一顶点的系数是零.

我们对一维链、二维链阐明了关系式③. 在拓扑学中可以证明关系式③对任意维数的胞腔和链都是成立的.

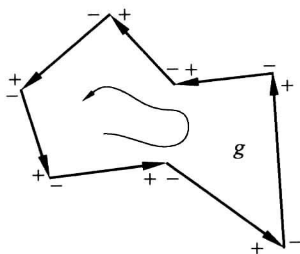


图 18.150

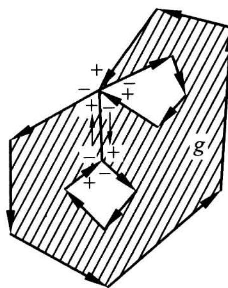


图 18.151

8.5　闭链与同调

边缘为零的链称为闭链,也称为圈. 后一称呼描述着这一事实,如果某些带有方向的线段顺次地首尾相接联成"圈"状(例如图 18.150 多边形的边),则形成边缘为零的链(即"圈").

由基本性质③可知任何链的边缘是闭链. 事实上,取链 $Z=\Delta x$ 是某一链 x 的边缘. 由③易知 $\Delta Z=\Delta(\Delta x)=0$,也就是 Z 是闭链(或圈). 对于由关系式 $Z=\Delta x$ 联系着的两个链 Z,x,可以说 x 是绷在圈 Z 上的胞膜. 例如多边形的内部是绷在它的边界上的胞膜. 因而,对每一链 Z,如果存在着绷在其上的胞膜(就是找出链 x 使 $Z=\Delta x$),则链 Z 是闭链(或圈). 然而,存在着这样的闭链(或圈)不可能绷上胞膜. 例如环面,如我们所知可以表示为由四个胞腔组成的剖分 T:一个零维、一个二维、两个一维胞腔(经圆和纬圆),并且所有关联系数都是零,也就是这些胞腔的边缘都是零. 由此可得,特别是剖分 T 的任何二维闭链的边缘都是零,因而剖分 T 的任何异于零的一维闭链上不可能绷上胞膜. 特别是,作为闭链(圈)的例的经圆或纬圆不可能绷上展布在环面上的胞膜.

这样,我们可以把闭链(圈)分成两种:

(1) 可以绷上胞膜的闭链(圈),即可以作为某些链的边缘的闭链.

(2) 不可能绷上胞膜的闭链(圈),即不能作为某些链的边缘的闭链.

我们称第(1)种的闭链为边缘链,第(2)种的闭链为非边缘链.

并且,如果两个非边缘链 Z_1,Z_2 的差 Z_1-Z_2 是边缘链(即可找到这样的链

x,使 $Z_1 - Z_2 = \Delta x$),则闭链 Z_1, Z_2 称为彼此同调.

现在我们把剖分 P 的所有闭链(圈)分类归并,两个同调的闭链算做属于同一类.我们得到某些同调类;属于同一类的闭链是彼此同调的,属于不同类的闭链彼此不同调.特别是,可以绷上胞膜的所有闭链组成一类,这是所有同调于零链的闭链类(因为关系 $Z = \Delta x$ 可以写成 $Z - 0 = \Delta x$),所以能绷上胞膜的闭链(边缘链)称为是同调于零的.

如果闭链 Z_1 同调于 Z'_1,Z_2 同调于 Z'_2(即 $Z_1 - Z'_1 = \Delta x_1, Z_2 - Z'_2 = \Delta x_2$),则闭链 $Z_1 + Z_2$ 同调于 $Z'_1 + Z'_2$

$$(Z_1 + Z_2) - (Z'_1 + Z'_2) = (Z_1 - Z'_1) + (Z_2 - Z'_2) = \Delta x_1 + \Delta x_2 = \Delta(x_1 + x_2)$$

这给出在同调类间可以进行相加:设 ξ_1, ξ_2 是某些同调类,而闭链 Z_1, Z_2 分别取自这两类,则我们取闭链 $Z_1 + Z_2$ 并把它的同调类表示为 $\xi_1 + \xi_2$.从上述加法可知如果在类 ξ_1 及 ξ_2 中分别取 Z'_1 及 Z'_2 来代替 Z_1 及 Z_2,则所得的和仍属于 $\xi_1 + \xi_2$.这样,以 ξ_1, ξ_2 所表示的同调类实际上仅依赖于类 ξ_1 及 ξ_2 本身,而并不依赖于 Z_1, Z_2 在这些类中的选取.

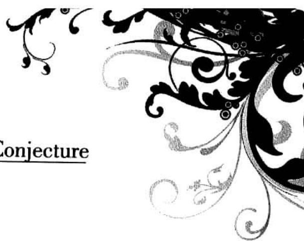
这样所定义的加法运算使所有同调类所成的集合形成群.这群的零元素是同调于零的闭链所属的类.我们称这群为所考察剖分 P 的同调群并表示为 $H_r(P)$;r 是所考察闭链的维数:当所考察的是零维闭链,我们得到零维同调群 $H_0(P)$,当所考察的是一维闭链,则得一维同调群 $H_1(P)$ 等.

我们以 $Z_r(P)$ 表示所有 r 维闭链的集合,以 $B_r(P)$ 表示所有 r 维边缘链(每一 r 维链可以作为某一$(r+1)$ 维链的边缘)的集合,则集合 $B_r(P)$ 包含在集合 $Z_r(P)$ 内.并且集合 $B_r(P)$,$Z_r(P)$ 本身都是群,因为两个闭链的和、差仍是闭链,两个边缘链的和、差仍是边缘链.换言之,$B_r(P)$ 是 $Z_r(P)$ 的子群.我们把 Z_1, Z_2 归并为同一同调类当且仅当差 $Z_1 - Z_2$ 属于子群 $B_r(P)$.所以同调类是商群 $Z_r(P)/B_r(P)$ 的元素;同调类的加法等同于在商群里所建立的加法.所以上面所阐述的同调群的定义可以归结为如下形式:r 维同调群 $H_r(P)$ 定义为闭链群 $Z_r(P)$ 关于它的子群——边缘链群 $B_r(P)$ 的商群,即

$$Z_r(P)/B_r(P) \tag{④}$$

同调理论的基本事实是下述关于"同调群是拓扑不变量"的(证明非常复杂的)定理.说得更清楚些:同调群 $H_r(P)$ 并不依赖于多面体的胞腔剖分的选择,而为多面体本身所确定;彼此同胚的多面体具有相同的(同构的)同调群.从而得出,特别地,在计算某些多面体的同调群时可以选取它的任意的胞腔分裂;自然,我们尽可能地去找较简单的剖分使同调群的计算较便利.





8.6　例

1. 计算球面 S^2 的同调群. 球面的最简单的胞腔剖分包含着两个胞腔: 一个顶点 a 及一个面 g. 这剖分不含一维胞腔, 所以一维链群 $L_1(S^2)$ 是平凡的(只有一个等于零的一维链). 从而作为一维链的边缘的零维链只能是一等于零的链. 但是所有零维链都是闭链(按定义, 任何零维链的边缘是零), 即 $Z_0(S^2) = L_0(S^2)$. 此外, 由于没有同调于零的闭链(除零链外), 所以同调群 $H_0(S^2)$ 与 $L_0(S^2)$ 相合, 这是由于 H_0 是群 $Z_0 = L_0$ 关于它的子群 B_0 的商群, 而 B_0 只有一个元素——零. 而每一个零维链具有形式 αa, 这里 α 是整数, 所以群 $H_0(S^2)$ 同构于整数加群(即自由循环群).

由于 $L_1(S^2) = 0$, 所有二维链都是二维闭链, 即 $Z_2(S^2) = L_2(S^2)$. 由于没有三维胞腔, 除零链外, 不存在同调于零的二维链, 故有 $H_2(S^2) = L_2(S^2)$ 且同构于整数加群. 由于群 $L_1(S^2)$ 是平凡的, 所以 $H_1(S^2)$ 也是平凡的.

我们用 Z 表示整数加群(自由循环群). 上述结果可以写作

$$H_0(S^2) = H_2(S^2) = Z, H_1(S^2) = 0 \qquad ⑤$$

(完全相同的证法, 可得 n 维球 S^n 的同调群为: $H_0(S^n) = H_n(S^n)$ 同构于整数加群; $H_r(S^n) = 0, 0 < r < n$. 按 $S^n: x_1^2 + x_2^2 + \cdots + x_{n+1}^2 = 1$, S^1 是平面上的圆周, S^2 是空间的球面)

2. 计算环面 P_1 的同调群. 利用考察过的关于环面的胞腔部分(一个顶点 a, 两个棱 r_1, r_2, 一个面 g), 我们得到任意的二维链具有形式 αg, 任意的一维链具有形式 $\beta_1 r_1 + \beta_2 r_2$, 任意的零维链具有形式 γa, 这里 $\alpha, \beta_1, \beta_2, \gamma$ 是整数, 由于所有关联系数都为零, 所以任何链的边缘为零. 从而所有链都是闭链, 而除零链外不存在同调于零的闭链, 故有

$$H_0(P_1) = L_0(P_1), H_1(P_1) = L_1(P_1)$$
$$H_2(P_2) = L_2(P_1)$$

从而得知 $H_0(P_1), H_2(P_1)$ 是自由循环群, 而 $H_1(P_1)$ 是由一对整数形成, 也就是两个群 Z 的直和(具有两个生成元的自由阿贝尔群). 上述结果可以写成

$$H_0(P_1) = H_2(P_1) = Z, H_1(P_1) = Z + Z \qquad ⑥$$

同样, 利用定向曲面 P_m(具有 m 个柄的球面)的胞腔剖分, 我们可得如下结果

$$H_0(P_m) = H_2(P_m) = Z, H_1(P_m) = Z + Z + \cdots + Z(2m \text{ 个加数}) \qquad ⑦$$

这里 $Z + Z + \cdots + Z$ 表示 $2m$ 个群 Z 的直和, 也就是具有 $2m$ 个生成元的自由阿贝尔群.

3. 最后, 我们考察射影平面 N_1 的同调群. 前面已看到存在着射影平面的胞

腔剖分:零维胞腔 a,一维胞腔 r,二维胞腔 g. 面 g 对棱 r 的关联系数是 ± 2,棱 r 对顶点 a 的关联系数是 0. 我们取这样的胞腔定向使 g 对 r 的关联系数是 $+2$. 这一胞腔剖分的零维、一维、二维链具有形式

$$x_0 = \alpha a, x_1 = \beta r, x_2 = \gamma g$$

这里 α, β, γ 是整数. 知道了关联系数,容易计算它们的边缘

$$\Delta x_0 = 0, \Delta x_1 = 0, \Delta x_2 = 2\gamma r \qquad ⑧$$

前面二式说明所有零维链都是闭链,及除零链外,不存在同调于零的零维闭链. 所以 $H_0(N_1) = Z$. 最后一式说明,如果 $x_2 \neq 0$(即 $\gamma \neq 0$),则 $\Delta x_2 \neq 0$. 也就是除零链外,所有二维链都不是闭链,所以 $H_2(N_1) = 0$. 余下的计算一维同调群,从上面的第二式知所有一维链都是闭链,而其中同调于零的(如上述第三式所示)闭链是在系数取偶数情况. 这样 $H_1(N_1)$ 是整数群 Z 关于它的偶数子群的商群,也就是模 2 的剩余类群 Z_2. 所以

$$H_0(N_1) = Z, H_1(N_1) = Z_2$$
$$H_2(N_1) = 0 \qquad ⑨$$

让读者去建立,模 m 的射影平面(用 Q_m 表示)具有如下的同调群

$$H_0(Q_m) = Z, H_1(Q_m) = Z_m$$
$$H_2(Q_m) = 0 \qquad ⑩$$

这里 Z_m 是模 m 的剩余类群.

请读者注意,在上面所考察的例中零维同调群都是自由循环群. 可以证明,更一般地,对于连通的多面体,零维同调群都是 Z.

如果多面体 P 不是连通的,而是由某些连通区组成的,则

$$H_0(P) = Z + Z + \cdots + Z$$

这里直和的加数的数目等于多面体 P 的连通区的数目 k(也就是 $H_0(P)$ 是具有 k 个生成元的自由阿贝尔群). 建议读者自己证明这一断言.

8.7 贝蒂数与欧拉示性数

在一般情况,同调群 $H_r(P)$(这里 P 是某一部分)具有形式

$$H_r(P) = Z + Z + \cdots + Z + K$$

这里 K 是有限群. 直和中 Z 的个数称为 r 维贝蒂(Betti)数①,记作 p_r.

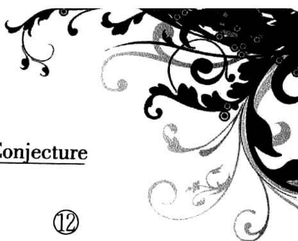
例如,定向曲面 P_m 的贝蒂数是

$$p_0 = 1, p_1 = 2m, p_2 = 1 \qquad ⑪$$

射影平面的贝蒂数是

① 贝蒂数这一术语为庞加莱所引进,以归功于意大利数学家贝蒂的卓越的发现.





$$p_0 = 1, p_1 = p_2 = 0 \quad ⑫$$

和同调群一样，贝蒂数也是拓扑不变量.特别是不依赖于胞腔剖分的选取.

我们考察过欧拉示性数这一重要的拓扑不变量.对二维剖分 P 的欧拉示性数由下式确定

$$\chi(P) = \Gamma - P + B$$

这里 Γ 是面数，P 是棱数，B 是顶点数.在一般情况，欧拉示性数由下式确定

$$\chi(P) = \sum_r (-1)^r \alpha_r^- \quad ⑬$$

这里 α_r^- 是剖分 P 的 r 维胞腔的个数(从而 $\alpha_0 = B, \alpha_1 = P, a_2 = \Gamma$).这里指出欧拉示性数可以借助于贝蒂数进行计算如下式

$$\chi(P) = \sum(-1)^r p_r$$

所示.这里 p_r 是多面体 P 的 r 维贝蒂数.由于贝蒂数是拓扑不变量，所以欧拉示性数也是拓扑不变量.由上式，我们容易计算曲面 P_m 及 N_1 的欧拉示性数如下

$$\chi(P_m) = 1 - 2m + 1 = 2 - 2m$$
$$\chi(N_1) = 1 - 0 + 0 = 1 \quad ⑭$$

9 同调理论的某些应用

如上所指出，欧拉示性数是一同调性概念.另一重要的同调性概念是“映象的度数”.本章的9.2、9.3将介绍这一概念极其有趣的几何应用.

9.1 曲面上的向量场

作为欧拉示性数的应用的重要例子是庞加莱关于曲面上的向量场的定理：考察任一定向曲面，它光滑地展布于空间，也就是在它的每一点上都具有切平面(不具有像多面体那样的顶点和棱).作为光滑曲面的例，有通常的球面、环面、双环面.现在提出如下问题：能否在所有曲面 P 上构造连续的方向场.也就是在曲面的每一点上，沿着曲面选取这样的方向(画在曲面上的箭头)，使当点变动时，箭头连续地变动？例如，如果在球面上的每一点画上自北向南的箭头(图18.152(a))，则除极点外，可得到连续的方向场，但在极点无法选取方向使与邻近点上的方向连续.如果在每一点上选取从西向东的方向(图18.152(b)).情况也是这样，所以在球面上构造连续的方向场遇到了困难.在环面上却易构造这样的方向场，如图18.153所示.

代替方向场可以考察切向量场.问题可以归结为如下形式：能否在曲面 P 的每一点上选取这样的非零向量，在这一点与曲面相切，使当点变动时切向量连续地变动？

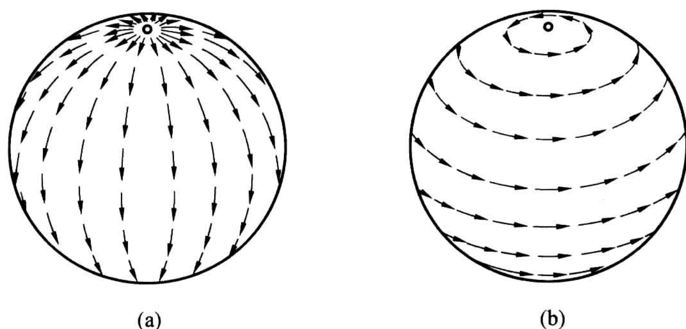


图 18.152

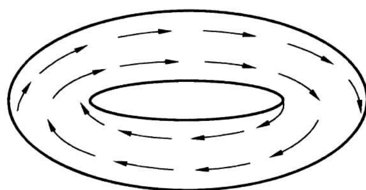


图 18.153

这一问题的解答是由庞加莱给出的:在曲面 P 上可以选取连续方向场(或连续的非零切向量场),当且仅当曲面 P 的欧拉示性数等于零.由公式 ⑭ 可知能够给出连续方向场的唯一定向闭曲面是 P_1,也就是环面.在球面上不可能给出连续方向场.这一事实可以另一形式归结为如下定理:如果在球面的每一点上可以装上一根"刺"(非零向量,不必与球面相切),使当点变动时,"刺"的方向连续地变动,则至少能找到一根刺它是垂直于球面的.事实上,如果所有的刺都不垂直于球面,那么这些刺在它们的切平面(刺所引出处的球的切平面)上的投影(图 18.154)形成了连续的非零切向量场,这是不可能的.

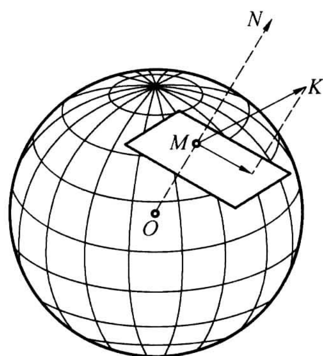


图 18.154

如上所述,在整个球面上构造连续方向场引起困难.如图 18.152 所示,除极点外,方向场是可以确定的.在极点处选取确定的方向是不可能的,因为在这一点的近傍有各种不同的方向.这种点称为方向场的奇点.在图 18.152 上除极点外无其他奇点.图 18.155 的(a),(b) 指出图 18.152(a),(b) 的方向场的北极附近的情况,而图 18.155(c) 描绘出另一种比较复杂的奇点(称为鞍点).我们考

察以奇点为中心的小圆周上的每一点处的向量的方向(图 18.156(a),(b),(c)).如果我们按逆时针方向环绕这圆周一周,可以看到,在情况(a)及(b),向量的方向也按逆时针方向环绕一周;而在情况(c)则按顺时针方向环绕一周.一般说,如果我们环绕以奇点为中心的小圆周一圈时,圆周上的向量环绕了 k 圈(在环绕方向相同时,k 是正值,相反情况取负值),则称 k 为所考察奇点的指数.在图 18.156(a),(b)情况的奇点的指数是 $+1$,(c)情况的奇点的指数是 -1.

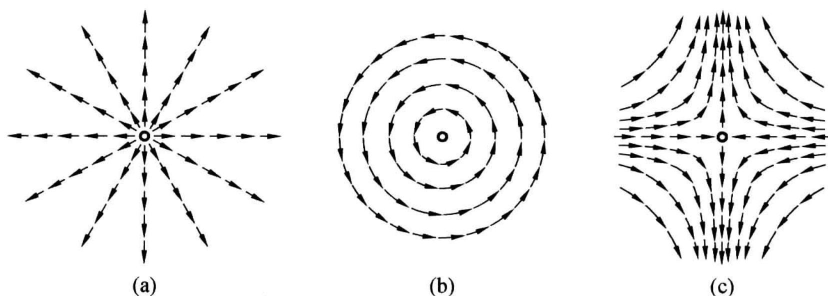


图 18.155

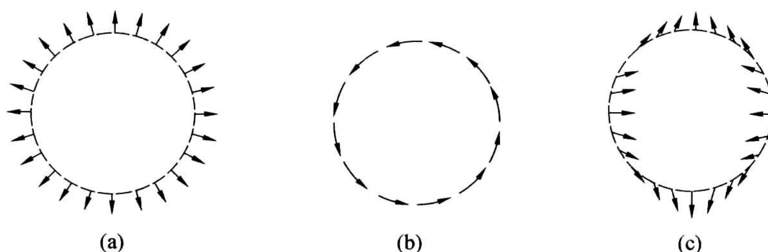


图 18.156

庞加莱定理的更完备的形式断言着:"如果在曲面 P 上给定了非零切向量场,除有限个奇点外到处连续,则这些奇点的指数的和等于曲面 P 的欧拉示性数."

对庞加莱定理作有趣而重要的推广者,在国外有霍普夫(Hopf)、斯蒂番尔(Stiefel)、斯汀劳特(Steenrod)等数学家.在国内(前苏联)有邦德里雅金和他的学生保斯脱尼高夫(Лостников)、罗赫林(Рохлин)、波尔金斯基等.

9.2 映象的度数与高斯 - 波内定理

设 P 与 Q 是两个定向曲面,f 是曲面 P 到曲面 Q 的连续映象.映象 f 并不假定是同胚的;可以想象映象 f 以种种方式把曲面 P 的几层折叠地落到曲面 Q 上.所谓曲面是定向的,是指对曲面上的每一点选定了环绕的方向(用环绕着这点的圆周上的箭头表示,并使曲面上邻近点处的方向一致(图 18.157)).在曲面 P

的某小片借助于映象f落到曲面Q的小片上时，如果这小片所对应的曲面P的（箭头）方向经映象f转为对应于曲面Q的（箭头）方向，则称所考察的P的这小片上的映象f是正的（图18.158(a)）；如果转为与曲面Q相反的（箭头）方向，则称映象f是负的（图18.158(b)）. 例如曲面P的某些小片通过映象f形成了摺层（图18.159），则在一部分的小片上的映象f是正的，另一部分的小片上的映象f是负的. 取曲面Q上某一点y及环绕着点y的小圆盘K（包含着点y）. 取曲面P上的所有小片，这些小片经映象f都映成圆盘K，并设每一小片到圆盘K上的映象f都是同胚的. 如果使（在小片上的）映象为正的小片的个数是p，使（在小片上的）映象为负的小片的个数是n，则称数$p-n$为点y附近映象f的度数. 换言之，映象的度数等于正（向）地覆盖点y的层数减去负（向）地覆盖点y的层数. 不难理解，曲面Q的任意点附近映象f的度数是一致的. 事实上，当点y在曲面Q上移动时，数p和n只有在通过摺层的边界时才会改变，而且在改变时，或者这两个数同时增加1，或者同时减少1，从而差$p-n$保持不变. 所以我们可以简

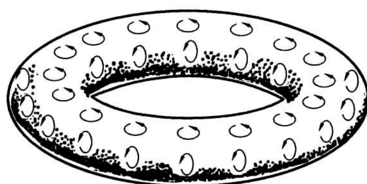


图 18.157

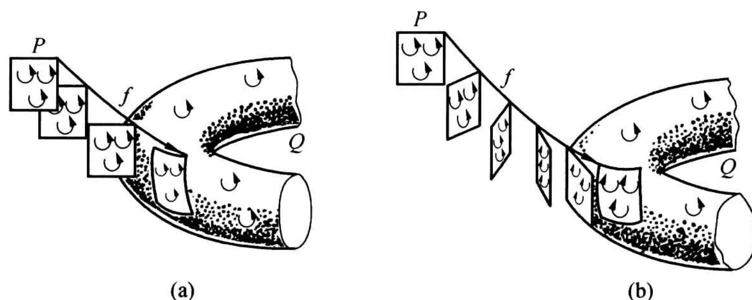


图 18.158

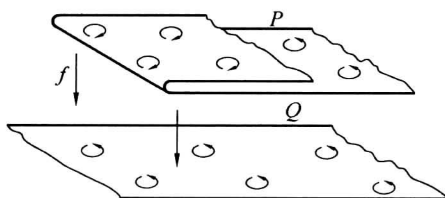


图 18.159

单地说映象的度数，不必指出是在曲面 Q 的哪一点附近计算度数的. 此外，注意，如果映象 f 连续地变化(形变)，则它的度数保持不变，也就是说新层的形成及层数的分布都保持不变(关于一个曲面到另一个曲面上的映象的度数，我们仅就最简单的直观的情况，除层外，映象没有其他"特异". 关于映象的度数的一般的严格的定义可以借助同调概念给出).

现在我们利用映象的度数这一概念去证明一个精美的公式，这是微分几何里熟知的所谓高斯(Gauss) - 波内(Bonnet)定理. 首先我们给出曲线在某一点的曲率的概念. 设 L 是某一曲线，x 是这曲线上的一点. 在曲线 L 上取包含着点 x 的一小段弧 $x'x''$，在点 x' 及 x'' 作曲线 L 的法线 $x'y'$ 及 $x''y''$(图 18.160). 设 α 是这两法线间的夹角，s 是弧 $x'x''$ 的长度. 作为"法线转动的平均速率"的比值 $\frac{\alpha}{s}$ 称为弧 $x'x''$ 上的平均曲率. 当包含点 x 的弧缩小趋向零时，相应的平均曲率所趋向的极限称为曲线 L 在点 x 处的曲率. 我们也可以把曲率叙述为如下形式. 取以某一点 O 为中心的单位圆 C，使曲线 L 上的点 x 与圆周上的点 $z = f(x)$ 建立如下的对应关系，使 Oz 平行于 L 上点 x 处的法线，这样得到了曲线 L 到圆周 C 上的映象 f. 设圆周上的点 z'，z'' 分别对应着 L 上的点 x'，x''，则弧 $z'z''$ 的长度在数值上等于角 $z'Oz''$ 的弧度数，也就是法线 $x'y'$ 与法线 $x''y''$ 间夹角 α 的弧度数(图 18.160)，所以比值 $\frac{\alpha}{s}$ 等于弧 $z'z''$ 的长与弧 $x'x''$ 的长的比. 当包含着点 x 的弧 $x'x''$ 的长度趋零时，弧 $z'z''(=f(x'x''))$ 的长度与弧 $x'x''$ 的长度的比趋向 L 上点 x 处的曲率.

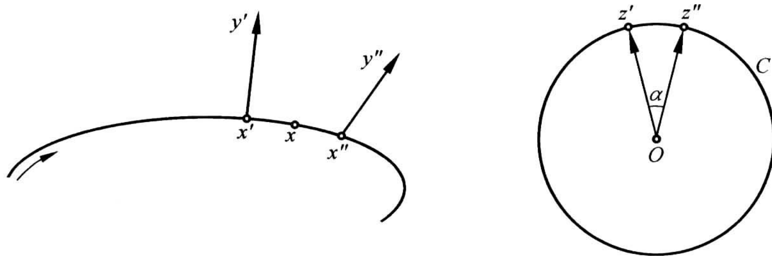


图 18.160

类似地可以定义曲面的曲率. 设 P 是某一曲面，为了定义在它的某一点 x 的曲率，我们考察从曲面 P 到单位球面的高斯球面映象. 取某一定向曲面 P，在它的点 x 上作一单位法线向量 $N(x)$(图 18.161). 取单位球面 S，在球心引出向量使平行且相等于 $N(x)$. 这向量的端点(显然落在球面上)用 $f(x)$ 表示. 这样得到的曲面 P 到单位球面 S 的映象(并且是连续的)称为高斯球面映象.

设 x 是曲面 P 上的某一点，M 是曲面的某一小片它包含着点 x. $f(M)$ 是小片 M 由球面映象 f 所得的象(在球面上). 用 $s(M)$ 及 $s(f(M))$ 表示 M 及 $f(M)$

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

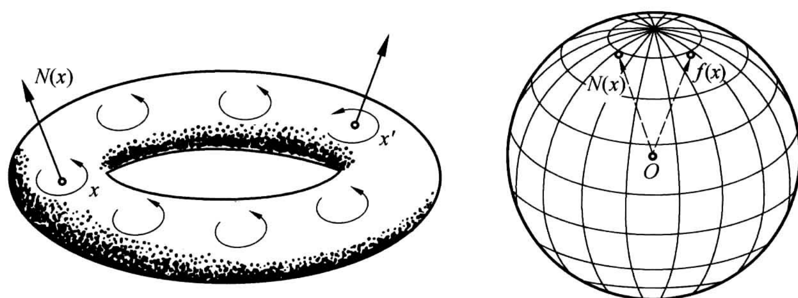


图 18.161

的面积,并且 $f(M)$ 的面积要带上符号:如果在片 M 上的映象 f 是正的,则片 $f(M)$ 的面积作为正的;如果 M 上的映象 f 是负的,则 $f(M)$ 的面积作为负的.比值 $\frac{s(f(M))}{s(M)}$ 称为曲面 P 上片 M 的平均曲率.当片 M 收缩于点 x,这比值所趋的极限称为曲面 P 上点 x 处的曲率(高斯曲率),记作 $k(x)$.我们可以有

$$s(M)k(x) \approx s(f(M))$$

当包含点 x 的小片 M 的面积愈小时,上式两端愈近似.

注意,如果曲面 P 的点 x 处的高斯曲率是正的,则在点 x 近傍曲面展布在这一点的切平面的同一侧,也就是凸的(图 18.162).曲面的这种点称为椭圆型的.如果曲率是负的,则点 x 近傍的曲面形成马鞍形的(图 18.163),与切平面相交于两直线,这样的点称为双曲型的.存在着抛物型的点,在这种点处高斯曲率等于零.

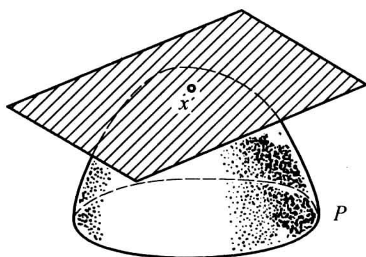


图 18.162

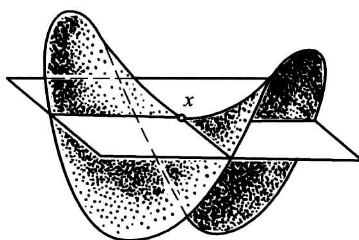


图 18.163

在曲面 P 的每一点 x 上定义了高斯曲率,也就是在曲面 P 上定义了函数 $k(x)$.这函数是连续的(设这曲面是由具有二阶连续导数的函数给出的).如果这曲面是闭的(就是没有边界的)及可以定向的,我们可以考察函数 $k(x)$ 在曲面 P 上的二重积分 $\iint_P k\mathrm{d}s$($\mathrm{d}s$ 是曲面的面积元).高斯 – 波内定理指出对于任何

定向闭曲面 P 具有公式①

$$\iint_P k\mathrm{d}s = 2\pi\chi$$

这里 χ 是曲面 P 的欧拉示性数.这公式表示着饶有意义的几何事实:虽然在每一点的高斯曲率本质地依赖于曲面在这一点近傍的结构(即弯曲的形式),然而高斯曲率在整个曲面上的积分不依赖于曲面的形式、弯曲情况以及在空间的分布等,而只依赖于这曲面的拓扑结构.

这里给出高斯 - 波内公式的另一解释,这解释联系着映象的度数的概念.为此,把曲面 P 分为小片 $M_1,M_2,\cdots,M_n$,在小片内分别取点 $x_1,x_2,\cdots,x_n$,则

$$\sum_{i=1}^{n} s(M_i)k(x_i) \approx \sum_{i=1}^{n} s(f(M_i)) \qquad (*)$$

上式左边是积分和,在小片 $M_1,M_2,\cdots,M_n$ 愈小时愈接近 $\iint k\mathrm{d}s$.右端是小片 $f(M_i)$ 的面积求和,这些小片覆盖球面的次数正好是映象 f 的度数(回忆面积 $s(f(M_i))$ 取正号,当 M_i 到 $f(M_i)$ 的映象是正的;取负号,在相反情况).这样式($*$)的右边部分等于 $\gamma\sigma$,这里 γ 是球面映象 f 的度数,σ 是球面 S 的面积.也就是右边等于 $4\pi\gamma$,故有

$$\iint_P k\mathrm{d}s = 4\pi\gamma$$

比较上述关系式与高斯 - 波内公式,为了证明后者只要证明 $\gamma = \frac{\chi}{2}$,也就是证明曲面 P 到球面 S 的球面映象的度数等于曲面 P 的欧拉示性数的一半.这一断言可以由庞加莱关于向量场的定理导出.设 $\boldsymbol{l}$ 是具有这样方向的向量,在曲面 P 上只有有限个点处的法线与 $\boldsymbol{l}$ 平行(具有这样性质的方向是存在的).在曲面 P 上的任意点 x 引向量 $\boldsymbol{l}$,把 $\boldsymbol{l}$ 在曲面上过点 x 处的切面上的投影记作 l_x.我们得到曲面上每一点 x 处的切向量 l_x.除去有限个点(这些点上的法线平行于 $\boldsymbol{l}$)外,l_x 在曲面上其他点处都异于零.这样,我们构造的连续切向量场仅有有限个奇点.在球面 S 上用同样方法构造向量场,这向量场正如图 18.152 所示具有两个奇点(在两极).显然,曲面上点 x 处的切向量 l_x 与球面上点 $f(x)$ 的切向量 $l_{f(x)}$ 互相平行.

设 $a_1,a_2,\cdots,a_q$ 曲面 P 上的向量场 l_x 的所有奇点,则 $f(a_i)(i = 1,2,\cdots,q)$ 也是球面上向量场的奇点(或者是球的南极,或者是北极).如果 $k(a_i) > 0$,则点 a_i 近旁的球面映象 f 是正的.所以当点 x 绕以 a_i 为中心的圆周作逆时针方向环行时,点 $f(x)$ 绕以 $f(a_i)$ 为中心的圆周也作逆时针方向环行.由于向量 l_x 与

① 这公式首先发表在波内 1948 年的著作中,其实早为高斯所发现.

向量 $l_{f(x)}$ 平行，所以当各自环绕圆周一周时，向量的方向作相同的变换. 这样，奇点 a_i 与 $f(a_i)$ 的指数相同，从而奇点 a_i 的指数是 $+1$. 如果 $k(a_i) < 0$，则点 a_i 近旁的球面映象 f 是负的. 所以当点 x 绕以 a_i 为中心的圆周作逆时针方向环行时，点 $f(x)$ 绕以 $f(a_i)$ 为中心的圆周作顺时针方向环行. 当做如上的绕行时向量 l_x 与 $l_{f(x)}$ 的方向仍作相同的变换. 从而，如果两个圆周（以 a_i 及 $f(a_i)$ 为中心的）都作逆时针方向绕行，则向量 l_x 的环绕圈数与向量 $l_{f(x)}$ 的环绕圈数差一符号. 所以，在这情况，奇点 a_i 与 $f(a_i)$ 的指数差一符号，从而奇点 a_i 的指数是 -1.

于是，点 $a_1, a_2, \cdots, a_q$——这是所有这样的点，它们由球面映象 f 映为球面上的极点，并且当它们近傍的映象 f 是正时它们的指数是 $+1$，当映象是负时，指数是 -1. 设在点 $a_1, a_2, \cdots, a_q$ 中，正向地映射到球的北极与南极的个数为 p' 与 p''，负向地映射到北极与南极的个数为 n' 与 n''，则这些奇点的指数的和等于

$$(p' + p'') - (n' + n'') = (p' - n') + (p'' - n'') = \gamma + \gamma = 2\gamma$$

（因为 $p' - n'$ 是球面上北极近旁映象的度数，$P'' - n''$ 是南极近旁映象的度数）. 但是由庞加莱定理，这些奇点的指数的和等于 χ，所以 $\chi = 2\gamma$. 到此完成了高斯－波内定理的证明.

9.3 代数学基本定理

借助于映象的度数的概念，可以给出代数学基本定理的精美的几何证明. 我们把复数 $z = x + iy$ 用平面上的点表示，x, y 作为这点的坐标（图 18.164）. 取球面 S，使它与平面相切于坐标原点，称这切点为球面的南极，相反方向的点 N 为北极（图 18.165），设 $z = x + iy$ 是任一复数 —— 也就是平面上的任一点. 线段 Nz 交

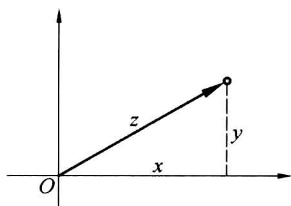


图 18.164

球面 S 于某一点，这点我们仍以 $z = x + iy$ 表示，作为复数 z 在球面上的象，从而所有复数在球面 S 上都有它们的象. 相反地，对球面上的点 a，容易看出它是平面上哪一个复数的象，可以引直线 Na，这直线与平面相交，给出所要求的复数. 但对球面上北极的那一点 N，不能给出所要求的复数. 我们规定点 N 所对应着的是一特异的“无穷”复数，用符号 ∞ 表示. 按上述规定，如果平面上的点无限地远离原点（按任何方向），则它在球面及 S 上的象就与北极靠近. 因而，球面 S 上的每一点映成某一复数，点 N 映成复数 ∞，球面上的其他点映成通常的“有限”复数. 球面 S 称为复数球面或黎曼球面. 注意，球面 S 是由平面添增一个无穷远点 ∞ 而得到的. 此外，添增某些点于平面可得到其他曲面 —— 射影平面.

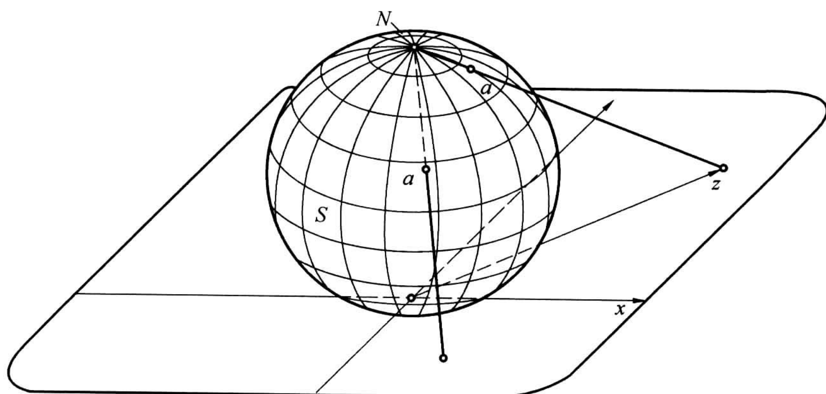


图 18.165

现在设 $f(z) = z^n + a_1z^{n-1} + \cdots + a_{n-1}z + a_n$ 是任意 n 次多项式.我们把 z 映到复数球面 S_1 上,把多项式的值 $f(z)$ 映到另一复数球面 S_2 上.球面 S_1 上每一“有限”点 $z = x + iy$ 对应着球面 S_2 的“有限”点 $f(z)$.并且如果点 z 靠近 ∞,则 $f(z)$ 也靠近 S_2 的 ∞ 点.实际上可把多项式 $f(z)$ 写成如下形式

$$f(z) = z^n(1 + \frac{a_1}{z} + \cdots + \frac{a_{n-1}}{z^{n-1}} + \frac{a_n}{z^n})$$

我们容易看到,当 $z \to \infty$(就是当 $|z|$ 无限增大),括弧内的表达式趋近 1,而因式 z^n 无限趋大.所以,可以补充定义 $f(\infty) = \infty$,从而得到整个复数球面 S_1 到复数球面 S_2 的连续映象.

为了确立方程

$$z^n + a_1z^{n-1} + \cdots + a_{n-1}z + a_n = 0$$

至少有一个解(如代数学基本定理所断言),可以去找出球面 S_1 上这样的点 z 使 $f(z) = 0$.也就是球面 S_2 上的点 0 至少是球面 S_1 上的一个点的象.如果不存在这样的点,则 S_2 上的点 0 不能为球面 S_1 的象 $f(S_1)$ 所覆盖,那么在球面 S_2 的点 0 的近旁映象 f 的度数等于零.由于映象 f 在任何点近旁的度数都是一样的,所以可以简单地说映象 f 的度数是零.为了证明代数学基本定理,只要证明映象 f 的度数异于零.我们指出这映象的度数等于 n(也就是与多项式 $f(z)$ 的次数符合,这是引进“映象度数”这一术语的合适理由).

让系数 $a_1, a_2, \cdots, a_n$ 趋零,则多项式 $f(z)$ 变动,S_1 到 S_2 的映象 f 作连续的形变(对 S_1 上的所有有限点 z 处,$f(z)$ 随着系数的变动而变动的连续性是明显的;对于无穷远点 ∞,由于多项式的首项系数总是 1,关系式 $f(\infty) = \infty$ 保持不变).最后,我们得到多项式 $f_1(z) = z^n$.但是,由于在连续形变时,映象的度数不变,所以由球面 S_1 到球面 S_2 上的映象 f 与 f_1 具有相同的度数.映象 f_1 的度数是很容易计算的.

考察平面上由原点 O 引出的 n 条射线构成的 n 个相等的角，每一角等于 $\frac{2\pi}{n}$(图 18.166)．当复数 z 自乘 n 次时，它的幅角增 n 倍，所以映象 f 把这些角中的每一个映射为整个球面 S_2，而且对角的内部说是同胚映象．所以通过映象 f_1，球面 S_1 的象 n 次覆盖球面 S_2，并且映象 f_1 处处是正的，所以映象 f_1 的度数是 n．

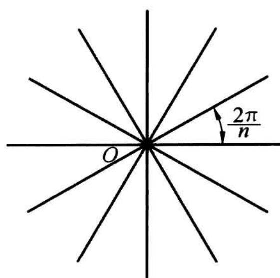


图 18.166



曲线是什么

第十九章

1 曲线概念的发展[①]

1.1 历史的概述

曲线是几何研究的主要对象之一.其原因首先是在于:曲线概念是在人类有关图形的绘制、地区疆界的确定、运动物体的轨迹的研究等实践活动中发生的.曲线的概念从实践中发生,并且在自然界和生产过程里诸现象的数学描述中,也有广泛的应用.

自古迄今,数学家们都注意着曲线的概念,其原因就在于此.科学家们会力求精密地定义,作为数学概念而言,曲线是什么?那就是要阐明:在实践中我们称之为曲线的那些事物里,什么为它们所共有?

这里企图直到近代才在前苏联数学家 П·С·乌利松(Урысон,1898—1924)的工作中获得成功,乌利松在 20 世纪的 20 年代中才给出曲线最一般的定义,有了这个定义才能彻底地研究这个概念的本质.但是 П·С·乌利松的工作之所以可能出现,是因为他深刻地批判地通晓了前人积累的一切重大科学资料的结果.因此要了解近代的曲线定义的自然性和必要性,我们就应该来注意这个概念怎样随着整个数学的发展而发展的.

① 原载自:巴尔霍民柯 A C.曲线是什么[M].华明,译.北京:科学出版社,1957.

欧几里得在他的著作《几何原本》中,定义曲线为无宽度的长(《几何原本》定义 2)或为表面的边界(定义 6)①这样的定义在一定程度上反映着曲线的特性,但在曲线概念的数学研究方面却毫无用处,因为它是由其他尚待定义的概念来定义的.正如人们所说:为了任何对象的数学研究,应该公理学地把它给定出来,那就是要指出这对象的一组性质,使它另外的性质可以逻辑地从这组性质推出来.

在当时科学发展的水平和实践所提出的要求的特性之下,欧几里得不可能在广阔的范围内来定义曲线的概念,因为他受上述一般成见所限制,在《几何原本》中,其注意力停滞于两种最简单而又是最常用的曲线——直线和圆周的研究上.

当然这并不是说:除直线和圆周以外,古人就不知道任何其他曲线.远在欧几里得以前,像狄诺斯特拉托斯(Dinostratus)的二次曲线就为人们所熟知了,过了一百年以后,阿波罗纽斯仔细地研究了圆锥曲线的理论,即平面和圆锥体侧面的交线——椭圆、双曲线和抛物线的理论(这些曲线的名称就是阿波罗纽斯的).在力学中也引起了阿基米德螺旋曲线的研究的必要性.然而这一切值是一些个别零散的事实,既没任何一般的曲线定义也没有它的研究方法.

在这一方面有决定性意义的一步,应归功于笛卡儿(R. Descartes, 1596—1650).原始资本积累时代的工商业的蓬勃发展促进了技术的迅速发展,也引起了自然科学,尤其是力学的空前发展.这种发展就需要数学工具,因为数学工具是力学定律的精密表示所必需的.在这数学工具的发展中,笛卡儿起了很大作用.

笛卡儿的理论在整个数学发展中,一般地具有重大的影响;尤其他的坐标方法,第一次使以当时算做很一般的形式来定义曲线成为可能.因此我们稍微仔细地来谈一下他的工作.

在平面上取定坐标系以后,则对应于平面上的每一点,我们可以确定一对实数为该点的坐标.于是不同的实数偶对应于不同的点,且每一实数偶对应着平面上完全确定的点,而此点以该实数偶为其坐标.这样就在平面上的点集和实数偶的集合之间,建立了所谓互相单值的对应.这个对应就可替每个曲线建立它的方程,也就是在它的点的坐标之间找到这样的一种关系式,对于这曲线上的所有点,这关系式皆成立,而对其他的任何点皆不成立,例如,以坐标原点为中心,以 r 为半径的圆周有方程

$$x^2 + y^2 = r^2$$

坐标轴的夹角的分角线有方程 $x - y = 0$ 等.

① 《Начала》Евклида, книги I—VI Гостехиздат, М.—Л. 1948,第 11 页.





替每个曲线建立它的方程之可能，给已知曲线的研究提供了非常一般而有力的方法；然而在曲线概念的一般定义的问题上，在没有以下列方式改换问题的提法以前，我们得不到任何新的东西.

设已知一方程，这方程有两个变数，将其各项移置等号之左，且以 $F(x,y)$ 表示方程左边的表达式(函数)，则可写成形式 $F(x,y)=0$，而且假设这方程的无穷个实数解，即有无穷个实数偶 x,y 适合此方程. 我们把 x 和 y 看做平面上的一点对应于某坐标系的坐标，其坐标适合方程 $F(x,y)=0$ 的平面上点的全体，则称为该方程所确定的曲线. 问题的这种提法的主要价值在于：我们现在可以给出曲线的一般的定义，而直到现在所知道的曲线的特例全部都包括在这个定义之内. 而且，一般地说，有多少不同的方程，就可以作出多少不同的曲线. 兹将此定义重述如下：凡坐标适合方程 $F(x,y)=0$ 的点的全体称为这方程所确定的曲线.

笛卡儿的发现对整个数学有决定性的意义，因为它一方面使几何对象可以用代数和分析的方法来研究，而另一方面也使几何的术语和方法可以用于代数和分析，从而使这些科目的研究大大地简单而明了了. 数学的辩证性就表现在数学发展中的两种对立倾向 —— 几何倾向和分析倾向之间互相渗透的影响中. 笛卡儿的曲线定义在当时算是非常一般的. 这个定义包括了所有代数曲线. 所谓代数曲线，就是代数方程所确定的曲线，即其方程的形式是 $F(x,y)=0$，而其中 $F(x,y)$ 是两个变数 x 和 y 的多项式. 多项式 $F(x,y)$ 的次数称为其代数曲线的阶. 因为所有平面上的直线在笛卡儿坐标中表示为一次方程的形式 $Ax+By+C=0$，而且所有这样的方程都表示直线，所以我们说：一阶代数曲线就是直线. 因为所有椭圆、抛物线、双曲线(或退化为二直线)都可以用二次方程来表示，而且所有二次方程，只要它有无穷个解，恒表示这些曲线之一，所以我们说：二阶的代数曲线是椭圆、抛物线或双曲线①. 这些事实就是解析几何学的内容，且基本上已为笛卡儿所通晓. 较高次的代数曲线的研究则是代数几何学的对象.

然而在那时已经知道：有些曲线根本不能用方程 $F(x,y)=0$ 来表示，这里 $F(x,y)$ 是指相当简单的函数，即有限个初等函数的组合；或即使可以用方程 $F(x,y)=0$ 来表示，但是，对于曲线的研究也丝毫无补. 那首先是动点轨迹的曲线. 就像阿基米德的螺旋曲线 —— 某点沿半直线均匀移动，而半直线又绕一定点以等角速度转动，该点所书的曲线. 如果以 r 表示动点离坐标原点的距离，而以 φ 表示转动着的半直线和笛卡儿坐标系统中 ox 轴正方向之间的夹角，那么对于阿基米德螺旋曲线，我们便得到 r 对 φ 的很简单的关系式

① 我们仅限于考虑实的曲线，方程没有实数解的情况，不在我们考虑的范围之内.

$$r = a\varphi$$

其中 a 是一个已知数.设 $M(x,y)$ 是阿基米德螺旋曲线上的任意一点,那么

$$r = \sqrt{x^2 + y^2}, \varphi = \arctan \frac{y}{x}$$

把这些数值 r 和 φ 代入等式 $r = a\varphi$,我们就得到阿基米德螺旋曲线在笛卡儿坐标中的方程

$$\sqrt{x^2 + y^2} - \arctan \frac{y}{x} = 0$$

然而这方程对于阿基米德螺旋曲线的研究却丝毫无补,因为对于 x 的每个数值,有无穷个 y 值与之相应;反之,对于 y 的每个数值,也有无穷个 x 值与之相应.

在动点的轨迹的研究方面,给出点依赖于时间的坐标是极其自然的.这就引入了所谓曲线的参数表达式,在这表达式中曲线的点的坐标被表示为某第三个变量 t(通常 t 指时间)的函数,t 称为参数

$$x = \varphi(t), y = \psi(t)$$

这样,如果动点 m 以等速度 v 沿一直线运动,而此直线通过原点且与 ox 轴的夹角为 α,那么动点的坐标依赖于时间而可用下面的公式表示

$$x = vt \cos \alpha, y = vt \sin \alpha$$

这两个等式便是直线的参数方程.

如果动点 m 均匀地沿着圆周运动,该圆以坐标原点 o 为中心,以 r 为半径,那么动点的坐标将依赖于时间,且可用下面的函数来表示

$$x = r \cos \omega t, y = r \sin \omega t$$

其中 ω 是半径 om 转动的角速度.这便是圆的参数方程.

阿基米德螺旋曲线的参数方程有以下的形式

$$x = vt \cos \omega t, y = vt \sin \omega t$$

这里 v 是动点沿半直线移动的速度,而 ω 是半直线绕坐标原点转动的角速度.

用参数方程来给出曲线,是完全合于对这概念所提出的一切要求的:大家所知道的一切曲线,包括代数曲线和超越曲线在内,都可以用这种方式来给定;且这种给定曲线的方式,最好地适应着作为动点的轨迹而得到曲线的主要方法.

定义曲线为动点的轨迹,并且用参数法给定大家所知道的曲线,这已经成为曲线概念新的推广的基础:如果平面上点的坐标 x 和 y 可以表示为某第三个变量 t 的函数(通常 t 看做时间,但也可以看做角度、弧长等),则这些点的全体将称为曲线.

当然在函数 φ 和 ψ 上加了一些限制,如果函数概念越广泛,则这些限制就越一般.由特例出发,在 19 世纪的后半叶,得到了下面的曲线的一般定义,法国





数学家若当(C.Jordan)以最清晰的形式界述了这样定义:如果平面上点的坐标是参数 t 在区间 $0 \leqslant t \leqslant 1$ 上所给定的连续函数①

$$x = \varphi(t), y = \psi(t)$$

那么这些点的全体就称为曲线.

然而立刻发现:若当的曲线定义又过分一般了:在 1890 年,意大利数学家皮亚诺(G.Peano)指出:可以在区间 $0 \leqslant t \leqslant 1$ 上选取这样的两个连续函数 $\varphi(t)$ 和 $\psi(t)$,使所有它的坐标适合方程:$x = \varphi(t), y = \psi(t)(0 \leqslant t \leqslant 1)$ 的点填满一整块正方形(包括其内点和边界点在内),那就是说,无论取正方形上的哪一个点 $m(x,y)$,总可以找到这样的参数值 $t(0 \leqslant t \leqslant 1)$,使

$$x = \varphi(t), y = \psi(t)$$

这个例子完全推翻了若当的曲线定义,如果就其一般地采纳这个定义:在这个定义的意义之下,“曲线”上的点将填满一整块面积,这与我们根据许多具体曲线的研究而形成的关于曲线的观念不能相容,因为我们所研究的那些具体的曲线总不能填满一整块面积.

1.2 皮亚诺“曲线”

为了要了解皮亚诺“曲线”是怎样做出来的,我们应该证明:线段 T 可以连续地映象到正方形 Q 上.我们更详细地来解释这是什么意思.假定我们已经有了任意线段 T 和任意正方形 Q.

问题就在于:对线段的每一点要找正方形某一点与之相应,而且正方形的每个点皆对应于线段的某个点.此外,对于我们有兴趣的线段映象应该是连续的,那就是说,这个映象应该合于下面的条件:只要把线段的点取得十分靠近,就可以使它相应的正方形的点任意靠近.

现在请大家来看,我们怎样来作这个映象.让我们把线段 T 分为四等份,自左至右依次以字母 $T_1^1, T_2^1, T_3^1, T_4^1$ 表示这些子线段.同样把正方形 Q 也分为四个相等的正方形,且以 $Q_1^1, Q_2^1, Q_3^1, Q_4^1$ 表示这些正方形(图 19.1).

线段 $T_1^1, T_2^1, T_3^1, T_4^1$ 以及与之相对应的正方形 $Q_1^1, Q_2^1, Q_3^1, Q_4^1$ 就叫做一级线段和一级正方形.进而,$T_1^1, T_2^1, T_3^1, T_4^1$ 中的每个线段再分为四等份,由 T_1^1 所分成的子线段依次用 $T_1^2, T_2^2, T_3^2, T_4^2$ 表示;用 T_2^1 所分成的子线段依次用 $T_5^2, T_6^2, T_7^2, T_8^2$ 表示;用 T_3^1 所分成的子线段用 $T_9^2, T_{10}^2, T_{11}^2, T_{12}^2$ 表示;由 T_4^1 所分成的子线段用 $T_{13}^2, T_{14}^2, T_{15}^2, T_{16}^2$ 表示;在第二次分割中所得到的16个线段 $T_1^2, \cdots, T_{16}^2$ 叫做二

---

① 回忆一下:如果在区间 $0 \leqslant t \leqslant 1$ 上给定的函数 $f(t)$ 具有下列的性质,就叫做连续函数:只要独立变数在区间 $0 \leqslant t \leqslant 1$ 中的任意二值相当靠近,就可以使得函数值相差任意小.

更精确一些,这意思应该这样来说:无论预先给定的正数 ε 多么小,总可以找到充分小的正数 δ,只要变元在区间 $0 \leqslant t \leqslant 1$ 中的两个值 t_1 和 t_2 相差小于 δ,则其相应的函数值相差小于 ε.

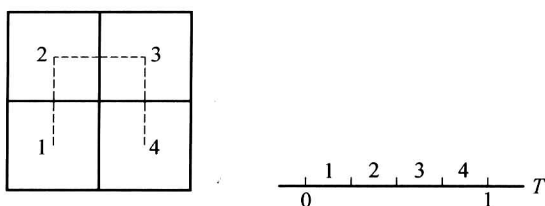


图 19.1

级线段.其次,再把 Q_1^1,Q_2^1,Q_3^1,Q_4^1 中的每个正方形分成四个相等的正方形.我们就得到 16 个二级的正方形.由 Q_1^1 所分成的正方形用 Q_1^2,Q_2^2,Q_3^2,Q_4^2 表示;由 Q_2^1 所分成的正方形用 Q_5^2,Q_6^2,Q_7^2,Q_8^2 表示;由 Q_3^1 所分成的正方形用 $Q_9^2,Q_{10}^2,Q_{11}^2,Q_{12}^2$ 表示;由 Q_4^1 所分成的正方形用 $Q_{13}^2,Q_{14}^2,Q_{15}^2,Q_{16}^2$ 表示.符号之选择致使其下标数相差 1 的二正方形至少有一个公共点.但因由同一个一级正方形所分成的四个正方形总有一个公共点 —— 这正方形的中心,所以在选择符号时只要注意:使每两个正方形 Q_4^2 和 Q_5^2,Q_8^2 和 Q_9^2,Q_{12}^2 和 Q_{13}^2 有公共点,那就是使一个一级正方形所分成的最后一个二级正方形和相邻的一级正方形所分成的第一个二级正方形有公共点(图 19.2).

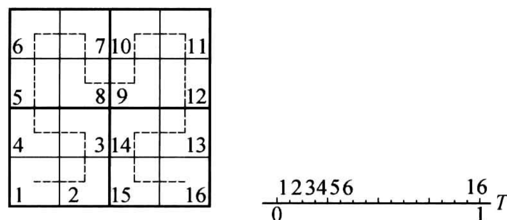


图 19.2

图 19.2 中的虚线说明:当自左至右顺序经过诸线段时,就按虚线所指示的次序通过诸正方形.

现在,再把每个三级线段分成四个相等的线段.这样一来,我们便得到 64 个三级线段,顺序(自左至右)以 $T_1^3,T_2^3,\cdots,T_{64}^3$ 表示;同样,把 16 个二级正方形的每一项再分为四个相等的正方形.我们就得到 64 个三级正方形,以 Q_1^3,$Q_2^3,\cdots,Q_{64}^3$ 来表示这些正方形,但此表示法致使 Q_1^3,Q_2^3,Q_3^3,Q_4^3 是由 Q_1^2 所分成;Q_5^3,Q_6^3,Q_7^3,Q_8^3 是由 Q_2^2 所分成,……

在选择符号时,也应该注意到:致使每相邻两三级正方形(下标数相差 1 者)有公共点.但因同一个二级正方形所分成的四个三级正方形都有一公共点(即这正方形的中心),故只需留意:使由某二级正方形所分成的最后一个三级正方形与由相邻的二级正方形所分成的第一个三级正方形有公共点(图 19.3).

图中的虚线表示:当自左至右顺序通过诸线段时,就按照虚线所表示的次

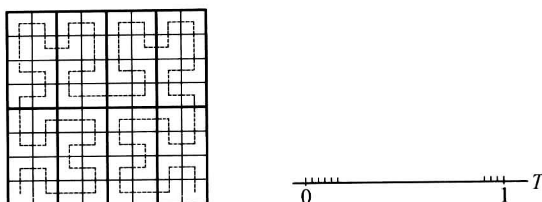


图 19.3

序通过诸正方形.

这种分割线段和正方形的手续可以无限制地继续进行;当分割的级数增加时,线段的长度和正方形的边长将同时趋于零,因为对于 n 级分割,线段的长度将等于 $1/4^n$,而正方形的边长将等于 $1/2^n$.现在对每个 n 级的线段 T_k^n,提供一个同级正方形 Q_k^n 与之对应,此正方形 Q_k^n 和指定的线段 T_k^n 有同一个号码 k.于是对于每级分割 n,在同级分割的线段与正方形之间就得到了一个互相单值的对应.在这些线段和正方形之间的对应,使我们能够作一个连续映,把线段 T 映象到正方形 Q 上.命 t_0 是线段的任一点.此点至少属于一个一级线段(不超过两个一级线段),也至少属于一个二级线段,也至少属于一个三级线段,循此渐近,便得到一系列线段,都包含着 t_0.对应于这些线段,便得到正方形的一个无穷序列.因为每个 n 级的正方形都包含在前一级的某个正方形里;又因为当级数 n 增大时,正方形的边长趋于零,故所有这些正方形有唯一的公共点 m_0,我们就命此点与点 t_0 对应.这样一来,我们就证明了:对于线段的每一个点对应着正方形的完全确定的点①.

现在需要说明:在这种对应之下,正方形的每个点至少和线段的一点对应.命 m_0 是正方形 Q 的任一点.此点至少属于一个一级正方形 Q_i^1,也至少属于一个二级正方形 Q_k^2,Q_k^2 包含在正方形 Q_i^1 内;也至少属于一个三级正方形 Q_l^3,Q_l^3 包含在 Q_k^2 内,循此继进.

我们来研究对应于这些正方形的线段.命这些线段是 T_i^1, T_k^2, T_l^3 等.每个线段都包含于前者,且当分割的级数增大时,线段之长趋于零.因此,所有这些线段有唯一的公共点 t_0.正方形 Q 的点 m_0 就和线段 T 的这一点 t_0 对应.

我们还要证明这种映象的连续性.如果 t_0 是线段的任意一点,则对于任意的 n 来说,这线段中所有十分靠近 t_0 的点 t,将和 t_0 在同一个 n 级的线段上,或在相邻的同级线段上.那么和线段 T 的两点 t_0 与 t 相对应的正方形的两点 m_0 与

① 当然,正方形的同一个点可能对应于线段的若干个不同的点,那就是说,我们的对应不是互相单值的.各级正方形的顶点就对应于线段的若干点:除了正方形 Q 的中心只对应于线段的三个不同的点以外,每个这样的顶点都对应于线段的四个不同的点.任何级正方形边上的点都对应于线段的若干个点,仅仅所有级正方形的内点才只与线段的一点对应.

m,将在同一个 n 级的正方形上,或在相邻的同级正方形上.所以只要把 t 点取得与 t_0 十分靠近,那么与之相对应的 m 点也就任意靠近于 m_0.这就证明了映象的连续性.

用我们所获得的线段到正方形的映象,就可以作皮亚诺"曲线".事实上,如果取数轴上的区间 $0 \leqslant t \leqslant 1$ 为线段 T,而把 XOY 平面上单位正方形取做 Q,它的点由不等式 $0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1$ 所决定.对于每一值 t,在已给的连续映象之下,乃得对应点 m 的坐标,于是这线段到正方形的映象就产生了两个连续函数 $x = \varphi(t), y = \psi(t)$.

函数 φ 和 ψ 和连续性可以由下列事实推出:如果把线段 T 的两点 t_0 和 t 取得充分靠近,则其所对应的点 m_0 和 m 亦将任意靠近,由此也就推出这两点的纵坐标及横坐标相差任意小.

这样便得到通过正方形的所有点的(用参数方程给出的)"曲线"的例子.

我们所考虑的线段至正方形的映象不是互相单值的;正方形上有无穷多个点存在,其中任意一点和线段上好几个(两个、三个或四个)点对应.我们可以证明:它与做成线段至正方形的连续映象的方法无关,在所有这样的映象之下,正方形总有这样的点存在,其每一点对应于线段的若干个点①.

1.3 单弧及由单弧所构成的曲线

可以将直线线段互相单值而连续地映象于其上的平面或空间的任一点集称为单弧.在平面上,如果取区间 $0 \leqslant t \leqslant 1$ 为直线线段,而取对应于已给的 t 值的点 M 的横坐标 x 和纵坐标 y 为函数 φ 和 ψ,则单弧恒可以用参数方程

$$x = \varphi(t), y = \psi(t), 0 \leqslant t \leqslant 1$$

来给出.在这种情况下,二函数 φ 和 ψ 都是连续的,并且对于二相异数值 t_1 和 t_2,下列不等式

$$\varphi(t_1) \neq \varphi(t_2) \text{ 或 } \psi(t_1) \neq \psi(t_2)$$

中至少有一个成立(当然也可能这两个不等式都成立).

我们所知道的曲线:如圆周、椭圆、双曲线、抛物线、正弦曲线、阿基米德螺旋曲线等,其中任意一个都是单弧的实例.

在空间,单弧可以由三个参数方程给出

$$x = \varphi(t), y = \psi(t), z = \chi(t)$$

而且对于参数的每两个数值 t_1 和 t_2,下列三个不等式

$$\varphi(t_1) \neq \varphi(t_2), \psi(t_1) \neq \psi(t_2), \chi(t_1) \neq \chi(t_2)$$

① 譬如,参考 Лузин Н.Н.,Теория функций действительного переменного(实变数函数论), Учпедгиз, Москва, 1940, 第 205 页.



至少有一个成立.螺旋曲线

$$x = r\cos t, y = r\sin t, z = at$$

的弧可以作为空间单弧的例子,其中 a 和 r 是常数,而 t 是参数,且 $0 \leqslant t \leqslant \infty$.

我们再引空间单弧的一个例子,这个例子给我们指出,"单曲线"也可以有多么奇怪的性质.基于这个目的,我们在皮亚诺曲线的方程

$$x = \varphi(t), y = \psi(t), 0 \leqslant t \leqslant 1$$

之外,再加入第三个方程

$$z = t$$

其坐标适合这三个方程的空间的点集合乃成一单弧,因为这个集合的点的三个坐标都与 t 连续相关,且不同的 t 值恒对应着集合中不同的点(因为第三个坐标不同)(图 19.4).

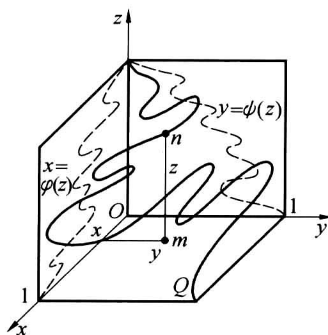


图 19.4

这曲线有一个特点:就是它在 xOy 平面上的投影是一个正方形.这样一来,在正方形的每一点的上面总有一个曲线的点,而在正方形的某些点的上面甚至有曲线的若干个(2,3 或 4) 点.

比方说:这个正方形的区域完全被一个屋顶所遮盖,而这个屋顶并不是由曲面而整个是由曲线所装制的,即屋顶本身是一支曲线.

作为圆周的互相单值且连续的影响的平面或空间的点集叫做单闭曲线.

在平面上,单闭曲线可以由下列的两个参数方程给出

$$x = \varphi(t), y = \psi(t), 0 \leqslant t \leqslant 1$$

且
$$\varphi(0) = \varphi(1), \psi(0) = \psi(1)$$

对于所有二数值 $t_1 \neq t_2$,只要其中有一个等于 0 和 1,则二不等式

$$\varphi(t_1) \neq \varphi(t_2), \psi(t_1) \neq \psi(t_2)$$

中至少有一个成立.

单闭曲线和单弧一样,也可能极其奇异地布于平面上.

可能在它上面的每一个点都没有切线;由单闭曲线所围成的区域的面积可能依赖于是否把其边界算在区域里.

但是单弧和单闭曲线是所有曲线中最简单的曲线,因为就其内在的性质而言,其结构是和直线线段及圆周一样;我们在上面谈到的其结构的各种特点是和它们在平面或空间的分布有关的.

任何点集,只要能分割为有穷支单弧,且这些单弧除其端点以外,互无其他

的公共点，那么这点集就可以称为曲线，我们所知道的一切曲线都合于这个定义. 例如圆周就可以看做是由二单弧 —— 其半圆所构成的曲线；贝努利(Bernoulli) 双叶线可以看做是由四支单弧所构成的曲线. 于是就得到一个充分广泛的曲线范畴，特别，所有代数曲线都包括在这个范畴之内. 但仍有一些曲线不能表示为有穷个单弧之和，这些单弧除其端点以外，互无其他的任何公共点. 譬如(图 19.5) 函数

$$y = \sin \frac{\pi}{2x}, 0 < x \leqslant 1$$

的圆形再加上其纵轴上的"极限线段"

$$x = 0, -1 \leqslant y \leqslant +1$$

这样所得到的曲线便是如此，我们在下面要证明，这个曲线不能表示为有穷个单弧的组合. 我们也看到：这个曲线不能表示为线段的连续映象(也就是说，它不是若当意义之下的曲线).

我们看到：如果在函数 $\varphi(t)$ 和 $\psi(t)$ 上，除连续性以外，不加任何其他的限制，那么若当曲线的范畴未免太广泛. 虽然如此，但是它并不包括在考虑力学和物理的许多问题时所遇到的所有曲线. 除已说明的曲线外，如圆周及填满于其中的螺旋曲线所组成的曲线也属于此(图 19.6). 在无线电学中研究振荡过程时必须考虑这样的曲线，而在考虑某些自动结构的稳定体系时，也必须考虑这样的曲线.

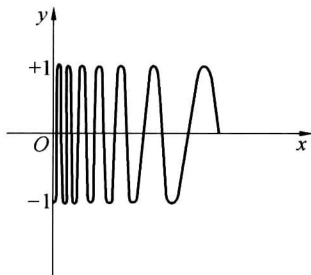


图 19.5

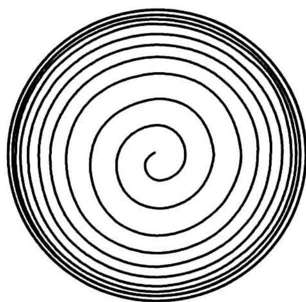


图 19.6

因此，直到现在，我们所考虑的定义曲线的一切办法都是有缺点的，因为每次都可以找到一些曲线不合于这些定义.

1.4 在定义曲线的问题中点集论的作用

到 19 世纪的末叶，在数学中，所谓集合论的观点开始越来越深入了，把所有的数学对象都看做某种元素的集合. 这里"集合" 概念本身不再被定义而看做基本的概念. 特别，几何图形被看做具有某种性质的点的集合. 例如，圆周就





是平面上与一定点等距离的诸点的集合(几何位置).把几何定理中的术语“几何位置”用“集合”来代替,我们最好在开始时就熟悉这个对整个近代数学是基本的概念.

康托在他的(19世纪70年代和80年代开始期间的)一系列的论文中,把集合论的观点很清楚地形成了.集合论的所有基本概念都归之于他.数学问题用集合论来论述是科学上一个重大的进展.特别,用集合论来论述,使康托把曲线概念的定义问题大大地向前推进了.我们开始来学习康托的曲线定义,以比较仔细地探求曲线概念.但在此以前,必须先引入集合论的一系列概念,因为没有这些概念,我们就不能作康托曲线定义,在学习乌利松的曲线的一般定义时,这些概念也还是必要的.一般地说,对于每一个人,不管他研究当代数学的哪一门,都有必要知道这些概念,因为点集论的基本概念在当代数学的整个邻域中都被采用着.

以后我们主要是让点集在所谓拓扑映象或同胚映象下不变的那些性质.

描述一下,我们可以说,拓扑映象就是把一个集合装置在另一个集合上,一个集合的任何两点不汇合于另一集合的一点,而且一集合的充分靠近的点对应于另一集合任意靠近的点,十分简略地说,拓扑映象把一个集合不重叠也不间断地转移到另一集合.

这种说法的精确的意思是:拓扑映象是互相单值互相连续的映象.

当我们读到单弧的概念时,就必须要牵涉拓扑映象.现在我们可以说:单弧是直线线段的拓扑影象.现在引几个各种同胚集合和非同胚集合的实例.

圆、正方形、三角形是彼此同胚的;圆柱体的侧面和圆环也是同胚的,所谓圆环,那就是夹于同心圆之间的点的集合(包括同心圆本身);全直线和无端点的区间也是同胚的.圆周、椭圆、正方形的周界等也是彼此同胚的.相反地,线段和正方形或线段和圆界,彼此都不是同胚的,线段和全直线也不是同胚的.(这里的线段是指闭区间,即带有端点的区间而言 —— 原译者)

可以证明:集合的某些性质,如连通性、列紧性、维数在同胚映象下保持不变,另一些性质,如两点之间的距离,点之是否属于同一直线或平面,这些性质在同胚映象下就不一定保留(这可以从上述的例子中看出).

研究集合在所有同胚映象下保持不变的性质的数学,就叫做拓扑学.这一门科学在20世纪的前半叶特别蓬勃地发展,而且引导近代数学的许多基本概念的建立,特别是曲线概念的建立.

苏联在拓扑学的发展中居于主导地位,这门科学的卓越的代表人物有:П·С·乌利松(Урысон),П·С·亚历山大洛夫(Александров)和Л·С·邦德里雅金(Понтрягин).

2 点集论中的一些知识

2.1 一般集合论的基本概念

在研究点集以前,我们回忆一下一般集合论的最基本的概念.

所有的集合都是由个别的元素所构成.如果 x 是集合 A 的元素,我们就说 x 属于 A,并且写

$$x \in A$$

如果两个集合 A 和 B 是由同一些元素所构成,也就是说,如果集合 A 的每个元素都属于集合 B,而集合 B 的每个元素也都属于集合 A,那么我们就认为这两个集合相等.

如果集合 A 和 B 相等,我们就写 $A = B$.

如果集合 A 的每个元素都属于集合 B,那么 A 就叫做 B 的子集.我们说 A 包含在 B 内,并且写 $A \subset B$;在这情况,我们也说 B 包含着 A,并且写 $B \supset A$.

如果 $A \subset B$ 且 $B \subset A$,那么 $A = B$.

集合 S 中的每一个元素至少属于集合 A 和 B 之一,则 S 称为二集合 A 和 B 的和集.集合 S 是集合 A 和 B 的和集的情况,我们写作

$$S = A \cup B$$

和集的概念可以推广到任意有穷的乃至无穷的集合族.

集合 P 的每个元素既属于集合 A 也属于集合 B,则 P 称为二集合 A 和 B 的交集.我们也说集合 P 是集合 A 和 B 的公共部分,并且写

$$P = A \cap B$$

交集的概念也可推广到任意的(有穷的或无穷的)集合族.

由集合 A 中不属于集合 B 的那些元素所构成的集合 D,叫做集合 A 和 B 的差集.由集合 A 减去集合 B 而得到集合 D 的情况,我们将写作

$$D = A \setminus B$$

在这里并没有假定集合 B 完全包含在集合 A 中.如果 $B \subset A$,那么

$$D = A \setminus B$$

就叫做集合 B 在集合 A 中的余集.在这种情况下,集合 B 和 D 称为互余.

有时为了论述上的一般性,必须纯粹形式地考虑一个元素也不包含的集合,即所谓空集.我们将用符号 0 来表示,按定义空集合是一切集合的子集.

例 1 如果 A 是区间 $0 \leqslant x \leqslant 2$ 的点的集合,而 B 是区间 $1 \leqslant x \leqslant 3$ 的点的集合,那么 $A \cup B$ 就是区间 $0 \leqslant x \leqslant 3$ 的点的集合;$A \cap B$ 是区间 $1 \leqslant x \leqslant 2$ 的点的集合;$A \setminus B$ 是半区间 $0 \leqslant x < 1$ 的点的集合.





集合的相加、相交和相减的运算具有数的加、乘和减诸运算所固有的许多性质,但集合的运算和数的运算也有许多相异之处.在这里我们不来研究集合上的运算的性质,因为在这本书中我们很少用到这些性质.这些问题的详细说明可以在集合论方面的专门书籍中找到①.

如果我们有了两个集合 A 和 B,指出集合 B 的什么元素对应着集合 A 的已给元素,那么我们就可以建立这两个集合的元素之间的一个对应.

集合 A 的元素 x 对应于集合 B 的元素 y 的情况,我们通常写作

$$y = f(x)$$

这里我们用符号 f 表示这样的规律,按照这个规律使集合 B 的元素来和集合 A 的元素对应.两个集合的对应的概念是普通的函数概念的推广.仅当变元 x 的数值所属的集合 A 以及函数值 y 所构成的集合 B 都是实数集合时才界定函数.

譬如,界定在区间 $-\frac{\pi}{2} < x < \frac{\pi}{2}$ 中的函数

$$y = \tan x$$

便建立了这区间中的数的集合和所有实数的集合之间的一个对应.

如果集合 A 的每个元素对应着集合 B 的某个完全确定的元素,那么我们就说:集合 A 映象到集合 B 中.在这里不必要集合 B 的每个元素都对应着集合 A 的某个元素.如果集合 B 的元素 y 对应于集合 A 的元素 x,则元素 y 叫做元素 x 的影象,而元素 x 叫做元素 y 的原象.如果集合 A 映象到集合 B 中,而 N 是集合 B 的子集,M 是集合 A 中所有其象属于集合 N 的元素构成的族,那么 M 叫做 N 的全原象.

如果集合 A 的每个元素都对应着集合 B 中某个完全确定的元素,并且集合 B 的每个元素都对应于集合 A 的某个元素,那么我们说:集合 A 映象到集合 B 上.这里并没有假定:集合 A 中不同的元素,对应着集合 B 中不同的元素.

如果集合 A 的每个元素都对应着集合 B 的某个完全确定的元素,并且集合 B 的每个元素都对应于集合 A 的某个元素,而且集合 A 的不同的元素对应着集合 B 中不同的元素,那么我们就说:集合 A 互相单值地映象到集合 B 上.在这种情况下,集合 B 的每个元素都对应着集合 A 中完全确定的元素;因此我们也可以说是集合 B 到集合 A 上的互相单值的映象.集合 B 到集合 A 上的这个映象称为原来的集合 A 到集合 B 上的映象的反映象,因此如果集合 A 互相单值地映象到集合 B 上,那么集合 B 也就互相单值地映象到集合 A 上.在这种情况下,也就说在集合 A 和 B 之间建立了互相单值的对应.

因而,在集合 A 和 B 的元素之间的互相单值的对应之下,集合 A 的每个元素对应着集合 B 的某个元素;集合 B 的每个元素对应着集合 A 的元素;集合 B

---

① 例如参考 П·С·亚历山洛罗夫(Александров):《集与函数的泛论初阶》(中译本)的第一章.

中不同的元素对应于集合 A 中不同的元素;集合 A 中不同的元素对应于集合 B 中不同的元素.

例 2 设 $abcd$ 是正方形,那么把它垂直地投影到 ab 边上,我们就得到正方形到直线 ab 内的映象.这个投影是正方形到线段 ab 上的映象.于是,线段 cd 被互相单值地映象到线段 ab 上.

有有穷集合也有无穷集合.如果可以把集合 A 的元素互相单值地对应于自然数列中从 1 到某个数 n 的数,则集合 A 叫做有穷集合.在这种情况下,我们说集合 A 包含 n 个元素.如果不可能建立这样的对应,那么集合就叫做无穷集合.

如果可以把集合 A 的元素互相单值地对应于所有自然数 $1,2,\cdots,n,\cdots$,那么集合 A 就叫做可列集.如果数 n 对应于集合 A 的元素 x,有时我就称 n 为元素 x 的号码,我们将把自然数 n 对应于元素 x 的情况写成 x_n;我们说集合 A 的元素用自然数标上了号码,而集合 A 和所有自然数的集合之间的这种对应将写成序列的列式

$$x_1, x_2, \cdots, x_n, \cdots$$

例 3 所有界于 0 和 1 之间的有理数的集合是可列的,所有有理数的集合也是可列的①.

如果可以把集合 A 互相单值地对应于所有实数的集合,那么就说集合 A 具有连续统的势.所有无理数的集合、直线上点的集合、从 0 到 1 的线段上的点的集合、平面上所有点的集合、空间所有点的集合等都可以作为具有连续统的势的集合的实例.

2.2 闭集和开集

在平面上或空间中取任意集合 R.这个集合 R 可以和全空间一致,但也可能仅构成空间的一部分.在我们现在所引入的一切定义中,值得注意的仅仅是集合 R 的点.至于空间的其余的点,只要它们不在集合 R 中,我们可以不考虑.

在直线、平面或空间,集合 R 的点所具有的一切性质中,对于我们最重要的是在集合 R 的两点 x 和 y 之间可以界定距离.

我们用 $\rho(x,y)$ 表示两点 x 和 y 之间的距离.在点与点之间的距离的许多性质中,只有下列的诸性质是我们需要的:

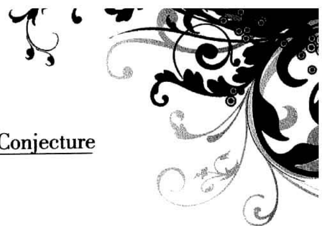
(1) 如果 x 和 y 是集合 R 的两个相异的点,那么其间之距离是正数.如果点 x 和点 y 重叠,那么其间之距离等于零.

(2) 对于集合 R 的任意两点 x 和 y

$$\rho(x,y) = \rho(y,x)$$

① 参考 П · С · 亚历山大洛夫:《集与函数的泛论初阶》(中译本)第 13 页,定理 8.





(3) 无论集合 R 的三点 x,y,z 怎样,下列不等式恒成立

$$\rho(x,z) \leqslant \rho(x,y) + \rho(y,z)$$

这个不等式我们可以称为三角形不等式.这个不等式实质上是这个几何定理——在一切三角形中每边小于其他两边之和——的记述.

按语 在以下的一切讨论中,不仅直线、平面或空间的点集合可以取做集合 R,一般地,可以采取任意元素的集合,只要这集合具有下面的性质:对于集合 R 的每一对元素 x 和 y 对应着数 $\rho(x,y)$,且适合条件(1),(2),(3).这样的集合称为度量空间,其元素称为点,而数 $\rho(x,y)$ 称为点 x 和点 y 间的距离.

n 维空间可以作为异于通常空间的点集的度量空间的实例.n 个数的有序组是这个空间的点,而两点 x 和 y 的距离定义如下:如果

$$x = (x_1,x_2,\cdots,x_n), y = (y_1,y_2,\cdots,y_n)$$

是 n 维空间的两点,那么

$$\rho(x,y) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \cdots + (y_n - x_n)^2}$$

这样引进的距离具有性质(1),(2),(3).我们不证明这个断言了①.n 维空间的一切子集都是度量空间.

在下面的全部说明中,都可以把 R 理解为任意的度量空间.如果读者感到过于抽象,也可以把 R 看做普通的欧几里得空间的子集,这在理解问题的实质方面是没有任何影响的.

集合的极限点的概念,对于整个点集论是很基本的.命 M 是集合 R 的一部分(子集)点所构成的集合.

x 是集合 R 的一点,如果在 x 的任意邻域中,总有集合 M 内的一点 y,y 异于点 x,那么点 x 就叫做集合 M 的极限点.“在任意邻域中”这句话应该这样来理解:不管正数 ε 多么小,在集合 M 中可以找到点 y,此点 y 和点 x 的距离小于 ε.点 x 可以属于集合 M,但也可以不属于 M.

命 R 是直线,于是对于点

$$1,\frac{1}{2},\frac{1}{3},\cdots,\frac{1}{n},\cdots$$

所构成的集合,零点为其极限点.对于直线上的有理点集合 A 而言,直线上一切的点(有理点及无理点)都是其极限点,而且有理点属于集合 A,而无理点不属于它.

F 是一个集合,如果所有点 x,只要它是集合 F 的极限点,都属于集合 F,那么 F 叫做闭集合(或闭集).

如果 R 仍然是直线,那么直线线段便是闭集,全直线也是闭集,而有理点

① 例如在 П · С · 亚历山大洛夫的《集与函数的泛论初阶》(中译本)中可以找到这个证明.

的集合对于全直线而言便不是闭集,因为直线上的无理点是有理点集合的极限点,但不属于这集合.

如果 F 是闭集,而 x 点不属于它,那么按定义,此点 x 就不是这集合的极限点;这意思是说:有这样的数 ε 存在,使集合 F 中没有一点与 x 点的距离小于 ε.

我们来介绍闭集合的两个基本性质.

定理 1 两个闭集合的和集是闭集.

事实上,如果 F_1 和 F_2 都是闭集合,而 x 是一个不属于和集 $F_1 \cup F_2$ 的点,则 x 就不属于集合 F_1 和 F_2 中的任一个.因为 F_1 和 F_2 都是闭集,那么就存在这样的数 ε_1,使集合 F_1 中没有一点与 x 点的距离小于 ε_1,也存在这样的数 ε_2,使集合 F_2 中没有一点与 x 点的距离小于 ε_2.如果假设 $\varepsilon_1 < \varepsilon_2$,那么无论在 F_1 和 F_2 中,都没有一点与 x 点的距离小于 ε_1,因此在集合 $F_1 \cup F_2$ 中没有一点与 x 点的距离小于 ε_1.因此,如果 x 点不属于集合 $F_1 \cup F_2$,那么它就不是这集合的极限点.

同样可以证明:任意有穷个闭集的和集仍是闭集.

定理 2 任意个闭集的交集仍然是闭集.

事实上,假定我们有了一族闭集合 $\{F_\alpha\}$.我们用 F 表示它们的公共部分,设 x 点不属于集合 F.这就是说,此点至少不属于集合族 $\{F_\alpha\}$ 中的某一个集合.假定这就是集合 F_{α_0}.因为此集合是闭集,故有这样的数 ε 存在,使集合 F_{α_0} 中没有一点与 x 点的距离小于 ε,更不待说,集合 F 中也没有一点与 x 点的距离小于 ε,因此,F 是一个闭集.

G 是一个集合,且 $G \subset R$,如果它的余集

$$F = R \setminus G$$

是一个闭集,那么 G 就称为开集合.由定理 1 和定理 2,就推出开集的下列的基本性质.

定理 3 两个(或任意有穷个)开集的交集是开集.

定理 4 任意个开集的和集是开集.

我们证明定理 3.命 G_1 和 G_2 是两个开集,G 是它们的交集.按开集的定义,集合

$$F_1 = R \setminus G_1$$

或

$$F_2 = R \setminus G_2$$

是闭集,按定理 1,其和集

$$F = F_1 \cup F_2$$

是闭集.

但是集合 $G = G_1 \cap G_2$ 是集合 F 的余集.事实上,如果 x 点属于集合 G,那么此点既属于集合 G_1,也属于集合 G_2,因此它既不属于 G_1 的补集 F_1,也不属



于 G_2 的补集 F_2,所以 x 点便属于集合

$$R \setminus (F_1 \cup F_2) = R \setminus F$$

反之,如果

$$x \in R \setminus F = R \setminus (F_1 \cup F_2)$$

那么 x 就既不属于集合 F_1,也不属于集合 F_2,因此 x 点又属于集合 G_1,也属于集合 G_2,即属于集合 G.

这样一来,我们就证明了集合

$$G = G_1 \cap G_2$$

是闭集

$$F = F_1 \cup F_2$$

的余集,因为 F 是闭集合,所以 G 便是开集.同样可以证明定理 4.

现在我们引入一个辅助概念,今后我们将不断地用到这概念.

命 x 是集合 R 的任一点,ε 是任一正数.集合 R 中一切与 x 点距离小于 ε 的点 y 的总体称为 x 点的 ε – 邻域.我们有时也称 ε – 邻域为球邻域,并且以 $S(x,\varepsilon)$ 表示.

如果 R 表示全空间,那么以 x 点为圆心,以 ε 为半径的球的所有内点的总体便是 x 点的 ε – 邻域.如果 R 是平面,那么 x 点的 ε – 邻域便是以 x 点为圆心且以 ε 为半径的圆的内部.

我们证明:x 点的 ε – 邻域是一个开集合[①].为此我们应该证明集合 $F = R \setminus S(x,\varepsilon)$ 是闭集合.事实上,命 y 是球邻域 $S(x,\varepsilon)$ 的任一点,则 y 点不属于集合 F.为了证明 F 是闭集,我们应该找到这样的数 δ,使所有和 y 点的距离小于 δ 的点 z,都不属于 F.只要假定

$$\delta = \varepsilon - \rho(x,y)$$

于是便有

$$\rho(x,z) \leqslant \rho(x,y) + \rho(y,z) < \rho(x,y) + \delta = \\ \rho(x,y) + [\varepsilon - \rho(x,y)] = \varepsilon$$

这样一来,如果点 $y \in S(x,\varepsilon)$,则所有十分靠近 y 的点 z 也属于 $S(x,\varepsilon)$.由此推出 $F = R \setminus S(x,\varepsilon)$ 是闭集,而 $S(x,\varepsilon)$ 是开集.

我们看到:点的球邻域是开集.另一方面,一切开集都可以表示成其点的球邻域的和集,且其加项可以是有穷个,也可以是无穷个.

事实上,命 x 是开集 G 的任一点;因为集合 $F = R \setminus G$ 是开集,那么就存在这样的数 ε,使一切与 x 距离小于 ε 的点 y 皆不属于 F,因此属于 G.但所有这样的点构成 x 点的 ε – 邻域,所有开集 G 的一切点都具有含于集合 G 中的球邻域

① 以后有时我们把 x 点的邻域理解为任意包含着 x 点的开集,而不仅是球邻域.

(当然这球邻域的半径依赖于此点).集合 G 便表示为这种邻域的和集的形式.

如果 R 是直线,那么 x 点的 ε - 邻域便是一个区间(无端点),其长度是 2ε,中心在 x 点.

因为所有开集都是其点的 ε - 邻域的和集,那么由此推出直线上的一切开集都是有穷个或无穷个区间的和集.可以作更精确的断言 —— 就是直线上的一切开集都是有穷个或无穷个互无公共点的区间的和集.这些区间的端点属于一个闭集,此闭集是已给开集的余集.

事实上,命 x 是开集 G 的任一点.按上述,就有一个包含着 x 点的区间存在,且此区间包含在集合 G 中.我们用 G_x 表示一切包含着 x 点且其自身却包含在集合 G 中的这些区间的和集,则 G_x 是包含着 x 点且含于集合 G 中的最大的区间.如果 y 是集合 G 的一点,不属于 G_x,那么与之相应的区间 G_y 和区间 G_x 就没有任何公共点,因为一旦区间 G_x 和 G_y 有一个公共点 z,则其和集 $G_x \cup G_y$ 也是包含着 x 点且含于集合 G 中的区间,并且此区间大于 G_x.

我们看到,集合 G 可以分解为互无公共点的区间的和集.这些区间集合至多是可列的.欲覆验此事,只要在每个区间中取某一个有理点.因为所有有理点的集合是可列的,那么其中属于所考虑的区间的那些有理点的集合或是有穷或为可列.

以后在考虑具体的例子时,在位于直线上的诸闭集中,我们必然遇到所谓康托完全集,此集合构成如下:由线段 $[0,1]$ 去掉区间 $\left(\frac{1}{3},\frac{2}{3}\right)$(当中的三分之一).再从剩下的两个一级线段 $\left[0,\frac{1}{3}\right]$ 和 $\left[\frac{2}{3},1\right]$ 的每个中去掉其当中的三分之一,即去掉区间 $\left(\frac{1}{9},\frac{2}{9}\right)$,$\left(\frac{7}{9},\frac{8}{9}\right)$.在这样得到的四个二级线段 $\left[0,\frac{1}{9}\right]$,$\left[\frac{2}{9},\frac{3}{9}\right]$,$\left[\frac{6}{9},\frac{7}{9}\right]$,$\left[\frac{8}{9},1\right]$ 的每个中,再去掉其当中的三分之一,却去掉区间 $\left(\frac{1}{27},\frac{2}{27}\right)$,$\left(\frac{7}{27},\frac{8}{27}\right)$,$\left(\frac{19}{27},\frac{20}{27}\right)$,$\left(\frac{25}{27},\frac{26}{27}\right)$.循此继进,以至一切自然数 n.

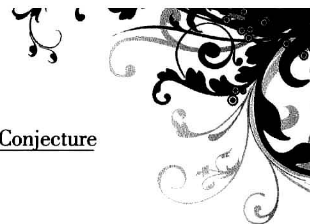
对于所有自然数 n,都施行这种手续,完成了这些手续以后,剩下的集合 P 就叫做康托完全集.

我们来注意康托完全集的一些性质.

(1) 集合 P 是闭集,因为它是一个下降的闭集序列的交集:这序列的第一个集合是由两个一级线段所构成,第二个集合是由四个二级线段构成,第三个集合是由八个三级线段构成,如此等等.

(2) 集合 P 不包含任何线段作为其子集.欲覆验此事,只要证明,无论集合 P 的点 x 和正数 ε 怎样,总可以找到一点 y,它与 x 点的距离小于 ε,且属于某个被挖去的区间.把 n 取得如此大,使





$$\frac{1}{3^n} < \varepsilon$$

因为每个 n 级线段的长度小于$\frac{1}{3^n}$,那么可以找到一点 y,它和 x 点的距离小于ε 而属于某个被挖去的n 级区间.

(3) 康托完全集在一个 n 级线段上的所有点的总体和康托完全集同胚.可以把线段[0,1] 相似地变换到 n 级线段上,其比例系数为$\frac{1}{3^n}$,以验明这个命题.

(4) 康托完全集具有连续统的势.

我们不来证明这个命题了.读者可在Π·C·亚历山大洛夫的《集与函数的泛论初阶》第 97 页找到这个证明.

由(3) 和(4) 立刻推出:

(5) 如果 x_0 是康托完全集的任一点,那么在这点的任何邻域中,都包含不可列个康托完全集的点.

运用球邻域的概念,我们来证明一个关于闭集和开集的定理.

定理5 如果A 和B 是两个无公共点的闭集,那么存在两个开集 G 和H,分别包含着集合 A 和B,而且也没有公共点.

证明 命 x 是集合A 的任一点.因为 B 是闭集而且x 点不属于集合B,所以 x 点就不是集合B 的极限点.因此存在这样的正数 ε_x,以至对于集合 B 的所有点 y

$$\rho(x,y) > \varepsilon_x$$

考虑球邻域

$$S\left(x,\frac{\varepsilon_x}{2}\right)$$

其中心在x 点,半径为$\frac{\varepsilon_x}{2}$.对于集合 A 的每个点x,我们都做这样的邻域,且以 G 表示一切这样的邻域的和集.

完全一样,对于集合 B 的每个点y 总可以找到这样的正数ε_y,以致对于集合 A 的每个点x

$$\rho(x,y) > \varepsilon_y$$

对于集合 B 的每个点y,做球邻域

$$S\left(y,\frac{\varepsilon_y}{2}\right)$$

其中心在 y 点,半径为$\frac{\varepsilon_y}{2}$.我们以 H 表示一切这样的邻域的和集.

集合 G 和H 是开集,因为它们每个都是球邻域的和集.现在我们来证明集合 G 和H 没有公共点.

假定相反,且命 z 是 G 和 H 的公共点.因为 $z\in G$,那么在集合 A 存在这样的点 x,使

$$z\in S\left(x,\frac{\varepsilon_x}{2}\right)$$

也就是
$$\rho(x,z)<\frac{\varepsilon_x}{2}$$

而因为 $z\in H$,那么在集合 B 中就存在这样的点 y,使

$$z\in S\left(y,\frac{\varepsilon_y}{2}\right)$$

也就是
$$\rho(y,z)<\frac{\varepsilon_y}{2}$$

但
$$\varepsilon_x<\rho(x,y)$$

同样
$$\varepsilon_y<\rho(y,x)=\rho(x,y)$$

因此

$$\rho(x,y)\leqslant\rho(x,z)+\rho(y,z)<\frac{\varepsilon_x}{2}+\frac{\varepsilon_y}{2}<$$
$$\frac{1}{2}\rho(x,y)+\frac{1}{2}\rho(x,y)=\rho(x,y)$$

由此得
$$\rho(x,y)<\rho(x,y)$$

所得到的矛盾证明了集合 G 和 H 没有公共点存在.

由(包含在集合 R 中的)集合 M 的点以及集合 M 的极限点所构成的集合 $\overline{M}$ 叫做集合 M 的闭包,由闭包的定义,直接推出 $M\subset\overline{M}$.

例 4 如果 M 是诸点

$$1,\frac{1}{2},\frac{1}{3},\cdots,\frac{1}{n}\cdots$$

的集合,那么把 0 点添到集合 M 中就得到 $\overline{M}$.如果 M 是直线 R 上的有理点的集合,那么 $\overline{M}=R$.如果 M 是区间 (a,b),那么其闭包 $\overline{M}$ 便是线段(或译为闭区间) $[a,b]$,此线段是由区间 (a,b) 添上其端点 a 和 b 而得到.

定理 6 任何集合 M 的闭包 $\overline{M}$ 是闭集.

证明 命 x 是集合 $\overline{M}$ 的极限点.我们应该证明 $x\in\overline{M}$.因为 x 是集合 $\overline{M}$ 的极限点,那么它的任意邻域 $S(x,\varepsilon)$ 总包含着 $\overline{M}$ 的点 y. y 点或者属于集合 M,或者是这集合的极限点.在这一种情况, y 点的任何 δ - 邻域总包含集合 M 的一点 z.如果我们把 δ 取得如此之小,以至 $S(y,\delta)\subset S(x,\varepsilon)$,那么 z 点将属于 $S(x,\varepsilon)$.所以在 x 点的任意邻域中恒包含集合 M 的某个点,于是集合 $\overline{M}$ 的极限点 x 也是集合 M 的极限点,即 $x\in\overline{M}$.

当我们研究乌利松的曲线定义时,下面的概念也是基本.

如果 G 是开集,它的所有不属于 G 的极限点的集合叫做集合 G 的边界.我



们用 $\Gamma p(G)$ 表示开集 G 的边界.

例如,假设 G 是直线 R 上的区间(a,b),那么 $\Gamma p(G)$ 是由这区间的端点 a 和 b 两点所构成.

所有开集的边界都是闭集.

事实上,按定义,开集 G 的边界是其闭包 $\overline{G}$ 中不属于集合 G 的那些元素的集合.因此

$$\Gamma p(G) = \overline{G} \setminus G$$

但是 $\overline{G} \setminus G = \overline{G} \cap (R \setminus G)$.因为集合 $\overline{G}$ 和 $R \setminus G$ 都是闭集,那么它们的交集

$$\overline{G} \cap (R \setminus G) = \overline{G} \setminus G = \Gamma p(G)$$

也是闭集.这就证明了我们的论断.

如果集合 M 的每两点之间的距离小于同一个数,则集合 M 称为有界集.

如果集合 M 的每两点之间的距离小于同一个数 ε,那么我们说,集合 M 的直径小于 ε,并且写

$$\delta(M) < \varepsilon$$

直线线段可以作为有界集合的实例.平面上的正方形、圆、三角形,空间的立方体、球、四面体等等都是有界集合.直线、平面和空间不是有界集.

按语 有界集合的每两点的距离的上界叫做此集合的直径①.线段的直径就是线段的长;正方形的直径就是它的对角线的长度;圆的直径就是它在初等几何的意义之下的直径.

为了以后叙述的方便,应当明确地指出,集合 M 的直径小于一个指定数是什么意思,而直径的定义本身在以后的说明中却用不到.

2.3 连通性

我们引进以下的概念,这个概念在整个点集论中具有重要意义而且在曲线的定义中也起着根本的作用.我们指的是集合的连通性的概念.

如果 M 是一个集合,把它任意分为两个子集 A 和 B,$M = A \cup B$,至少在这两个子集的一个中,可以找到一点是另一子集的极限点,那么集合 M 就称为连通集.因此,如果集合 M 不是连通集,那么可以把它分为两个非空子集 A 和 B,这两子集不仅没有公共点,而且还具有以下的性质:这两个集合中没有一点是另一集合的极限点,即有这样的分割 $M = A \cup B$ 存在,对此分割

$$(\overline{A} \cap B) \cup (\overline{B} \cap A) = 0$$

我们所考虑的问题,都在集合 R 中进行,让我们在集合 M 和集合 R 一致的特殊情况下来考虑连通性的概念.如果我们假定集合 R 不是连通集,且 $R =$

① 设 M 是一个实数集合,如果这集合的所有数都小于同一个数 L,则此集合 M 称为囿于上.具有下列性质的数,即集合 M 的一切数皆不超过这个数,这样的数中最小的称为集合 M 的上界.

$A \cup B$ 是集合 R 的一个分割，A 和 B 不是空集，并且其中没有一个集合包含另一集合的极限点．那么由此推出：集合 A 和 B 是 R 中的闭集．但因 A 和 B 互为余集，那么由此推出，集合 A 和 B 也是 R 中的开集．因此，整个集合 R 为连通集，当而且只当集合 R 不能表示为两个不相交的非空子集合之和，此二子集的每个是 R 中开集同时也是 R 中的闭集．

我们证明：直线线段是连通集．事实上，命 $[a,b]$ 是数直线上的线段．和定理的说法相反，我们假定线段上点的集合不是连通集，命 F_1 和 F_2 是两个无公共点的非空间集，且其和集就是线段 $[a,b]$．假定 b 点属于集合 F_2；那么集合 F_1 便是由线段上的 $x(x < b)$ 点所构成，用 c 表示集合 F_1 的点的上确界．c 点或者属于集合 F_1，或者是这集合的极限点；然而因为集合 F_1 是闭集，那么在这种情况下，$c \in F_1$．如果 c 点不是 b 点，那么所有点 $x > c$ 便属于集合 F_2．因此 c 点是这个集合 F_2 的极限点，所以属于 F_2．如果 c 点和 b 点一致，那么按照假定，它也属于集合 F_2，所以在一切情况下，$c \in F_2$，因此我们就找到了集合 F_1 和 F_2 的公共点 c，但我们假定了 F_1 和 F_2 不相交．所得到的矛盾证明了线段 $[a,b]$ 的连通性．

直线上所有有理点的集合可以作为非连通集的实例．事实上，如果我们以 A 表示所有小于 $\sqrt{2}$ 的有理数的集合，而以 B 表示所有大于 $\sqrt{2}$ 的有理数的集合，那么集合 A 没有一点是集合的极限点，且集合 B 也没有一点是集合 A 的极限点．

由一点所构成的集合以及空集都被认为是连通集．

我们来证明一系列关于集合的连通性的定理，这些定理在以后将不断地用到．

定理 1　如果 A 和 B 是 R 中的两个无公共点的闭集或开集，而且连通集 M 包含在这两个集合 A 和 B 的和集中，即

$$M \subset A \cup B$$

那么集合 M 就必定包含在这两个集合 A 和 B 的一个之中，即 $M \subset A$ 或 $M \subset B$．

证明　命 $M \subset A \cup B$．于是 $M = (M \cap A) \cup (M \cap B)$，因为集合 A 和 B 或者同时是 R 中的开集，或者同时是 R 中的开集，而且没有公共点，所以集合 $M \cap A$ 和 $M \cap B$ 中没有一个包含着另一集合的极限点．而因为 M 是连通集，所以这两集合中的一个，例如 $M \cap B$，应该是空集，由此推出 $M \subset A$．

我们常常依赖于这被证明了的定理的这样的推论．

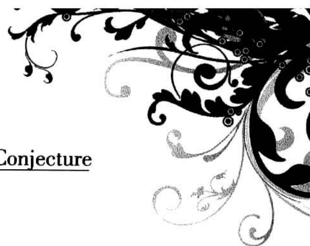
定理 2　如果连通集 M 与开集 G 以及闭集

$$F = R \setminus G$$

皆有公共点，这里 G 和 F 是 R 中互余的集合，那么集合 M 就与集合 G 的边界相交．

事实上，因为





$$\Gamma p(G) = \overline{G} \setminus G$$

所以整个集合 R 就可以表示为下列的三个互无公共点的集合的和集的形式：集合 G，集合 G 的边界及 G 的闭包的余集，即

$$R = G \cup \Gamma p(G) \cup (R \setminus \overline{G})$$

如果我们假定，连通集 M 不与集合 G 的边界相交，那么由此将推出：M 包含在无公共点的开集 G 和 $R \setminus \overline{G}$ 的和集中，而根据上面的定理也就推出：M 或完全包含在 G 中，或完全包含在 $R \setminus \overline{G}$ 中，这和本定理的条件不合.

定理3 如果集合 R 的任意两点 x,y 都属于某个连通集 C_{xy}（当然这集合与点 x,y 有关），那么集合 R 便是连通集.

事实上，如果集合 R 是非连通集，那么它将可以表示为两个无公共点的非空闭集 F_1 和 F_2 的和集. 在 F_1 中取一点 x，在 F_2 中取一点 y，我们可以用连通集 C_{xy} 连接这两点. 因为 F_1 和 F_2 是无公共点的闭集，所以按定理1，集合 C_{xy} 应该完全包含在 F_1 中或完全包含在 F_2 中. 所得到的矛盾证明了集合 R 是连通集.

以上我们证明了直线线段是连通集. 因为直线上的任意两点总可以考虑为此直线上的一个线段的端点，所以根据上面的定理可以推出：全直线是连通集. 一般在直线上仅仅有下列诸连通集：线段 $[a,b]$，区间 (a,b)，半区间 $[a,b)$ 或 $(a,b]$，全直线，半直线（无穷的半区间），具有一个出发点的半线. 这个论断的证明是根据以下的事实：如果两点属于这直线上的一个连通集，那么以这两点为端点的线段便属于这连通集.

因为平面上或空间中的任意两点都可以用线段来连接，所以由此推出：平面和空间都是连通集. 一般地说，所以凸集合都是连通集. 所谓凸集合就是这样的集合：其中任意两点都可以用全部属于这个集合的线段来连接. 特别，如平面上的三角形、正方形、圆，空间中的四面体、立方体、球等都是凸集合.

定理4 如果两个连通集 A 和 B 至少有一个公共点 x，那么其和集 $S = A \cup B$ 也是连通集.

事实上，如果假定集合 S 不是连通集，那么它就可以表示为两个互无公共点的非空闭集 F_1 和 F_2 的和集形式

$$S = F_1 \cup F_2$$

我们假定集合 A 和 B 的公共点 x 属于集合 F_1；于是根据定理1，集合 A 和集合 B 都包含在 F_1 中，所以集合 F_2 是空集，即 $S = A \cup B$ 连通集.

无论是这定理本身，或是这定理的证明，都可以几乎逐字地推广到任意多个集合的和集的情形，也就是以下的定理成立.

定理4′ 任意一组连通集，只要它们至少有一个公共点 x，那么这组集合的和集便是连通集.

因为每个线段都是连通集，所以所有折线也是连通集，所谓折线乃是一个线段序列，其每相邻两线段具有公共端点．一般地说，一个连通集的有穷序列，如果其中每相邻两集合有公共点，我们称这样的序列为集合链．于是构成集合链的诸连通集的和集也是连通集．

定理 5　集合 C 是包含在 R 中的连通集，如果把它的任意个极限点添到集合 C 中，那么所得到的仍然是连通集 C_0．特别，如果集合 C 是连通集，那么它的闭包 $\overline{C}$ 也是连通集．

事实上，命集合 C_0 是由连通集 C 添进它的某些极限点而得到的．假定集合 C_0 表示为两个无公共点的非空闭集（对集合 C_0 而言）F_1 和 F_2 的和集．因为集合 C 是连通集，那么按定理 1，它就应该完全包含在这两个集合的一个当中．譬如命 $C \subset F_1$．如果 x 是集合 C_0 的任意一点，那么它是集合 C 的极限，更不待说，它是集合 F_1 的极限，因为 F_1 包含着 C．但是因为 F_1 是 C_0 中的闭集，则 x 点属于集合 F_1．因此集合 C_0 的一切点都属于 F_1，所以集合 F_2 是空集，于是 C_0 是连通集．

在我们以后的说明中，局部连通性的概念起着极根本的作用．

如果集合 R 的每个点都有任意小的连通邻域，那么集合 R 叫做局部连通集．这就是说，无论集合 R 的 x 点及正数 ε 如何，总存在着连通开集 V 包含着 x 点，且其直径小于 ε；或者说，无论包含着 x 点的开集 U 如何，总存在一个连通开集 V 包含着 x 点且包含在 U 中．

直线、平面、空间都是局部连通集，因为直线、平面或空间的任意点的一切球邻域都是连通集．直线上坐标为

$$1,\frac{1}{2},\frac{1}{3},\cdots,\frac{1}{n},\cdots,0$$

的点的集合可以作为非局部连通集的实例．因为包含着 0 点的一切开集皆包含着无穷个形状为 $\frac{1}{n}$ 的点，故非连通集．所以该集合不是局部连通集．

从前提到过的函数

$$y = \sin\frac{1}{x}, 0 < x \leqslant 1$$

的图形上的点及极限线段

$$x = 0, -1 \leqslant y \leqslant 1$$

的点所构成的集合，这集合就可以作为连通集而非局部连通集的实例．

由互相更换的一序列垂直线段和水平线段

$A_1: x = \frac{1}{2}, -1 \leqslant y \leqslant 1$

$B_1: \frac{1}{4} \leqslant x \leqslant \frac{1}{2}, y = -1$



$A_2: x = \frac{1}{4}, -1 \leqslant y \leqslant 1$

$B_2: \frac{1}{8} \leqslant x \leqslant \frac{1}{4}, y = +1$

$A_3: x = \frac{1}{8}, -1 \leqslant y \leqslant 1$

$B_3: \frac{1}{16} \leqslant x \leqslant \frac{1}{8}, y = -1$

$\vdots$

$A_n: x = \frac{1}{2^n}, -1 \leqslant y \leqslant 1$

$B_n: \frac{1}{2^{n+1}} \leqslant x \leqslant \frac{1}{2^n}, y = (-1)^n$

$\vdots$

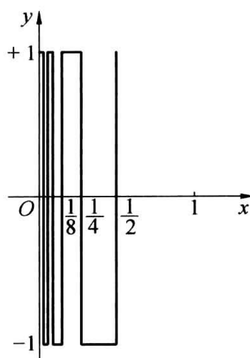


图 19.7

和极限线段

$$A_0: x = 0, -1 \leqslant y \leqslant 1$$

所构成的集合(图 19.7) 和上述集合同胚,因为我们可以用它来代替上述集合,考虑起来就比较方便.而这集合是由无穷阶的折线线段所构成的连通集合的闭包,故为连通集.然而它在极限线段中的任一点都不是局部连通的,因为如果取此点充分小的邻域,那么它将包含无穷个垂直线段的点,而与连接着它们的水平线段没有公共点,因此这邻域不是连通集.

2.4 列紧性

对于整个点集论而言,列紧性概念是很重要的.

R 是一个集合,如果它的一切无穷子集 M 在 R 中都有极限点,那么集合 R 就叫做列紧的集合.列紧的集合常常简称为列紧集.

我们证明:直线线段是列紧集.

事实上,命 M 是线段 ab 上的任意无穷集.我们用 c 点把线段 ab 分成两半.于是至少其中有一个线段 ac 和 cb 仍包含集合 M 中的无穷个点.再把这线段分成两半,在新得到的诸线段之一的线段上有 M 的无穷个点.进而作同样的推论,我们便得到一个下降的线段序列

$$T_0 \supset T_1 \supset T_2 \supset T_3 \supset \cdots \supset T_n \supset \cdots$$

其中每个线段都较前一线段短一半,而且包含着集合 M 中的无穷个点.

按照关于收缩线段的定理①,存在着唯一的点 x_0,属于一切线段 T_n.因为无论任意包含着 x_0 点的区间,总可以找到充分大的号码 n,使得具有这号码的

① 在 Π · C · 亚历山大洛夫的《集与函数的泛论初阶》(中译本) 第 28 页中,有本定理的证明.

线段 T_n 完全包含在取定的区间之内,于是这区间将包含集合 M 的无穷个点.这就证明了 x_0 点是集合 M 的极限点,因为在 x_0 点的任意邻域中包含着集合 M 异于 x_0 的点.

同样可以证明:正方形是列紧集.

事实上,命 Q_0 是一正方形,其边平行于坐标轴,命 M 是这个正方形的点的一个无穷集合.用平行于坐标轴的直线把正方形 Q_0 分成四个相等的正方形,并且用 Q_1 表示包含着 M 的无穷个点的任一正方形.再把 Q_1 分成四个相等的正方形,用 Q_2 表示包含着 M 的无穷个点的任一正方形.用同样的办法,进而我们得到正方形的一个降序列

$$Q_0 \supset Q_1 \supset Q_2 \supset \cdots \supset Q_n \supset \cdots$$

其中每个正方形都包含着 M 的无穷个点,而且其中每个后面的正方形的边长比前面一个正方形的边长小一半.

如果我们把这些正方形投影到 ox 轴和 oy 轴上,那么在每个轴上就得到一个收缩的线段序列

$$T_0^x \supset T_1^x \supset T_2^x \supset \cdots \supset T_n^x \supset \cdots$$

和

$$T_0^y \supset T_1^y \supset T_2^y \supset \cdots \supset T_n^y \supset \cdots$$

其边长随 n 的增大而趋于零.

我们用 x_0 表示所有线段 T_n^x 的公共点,用 y_0 表示所有线段 T_n^y 的公共点,并且考虑平面上具有坐标 x_0 和 y_0 的点 m_0.点 m_0 属于一切正方形 Q_n,故为集合 M 的极限点,因为无论任意以 m_0 为中心的圆,总可以找到这样的号码 n,使正方形 Q_n 完全包含在这个圆中;而因为每个正方形 Q_n 都包含 M 的无穷个点,因此有 M 的无穷个点落在这个圆中.这就证明了 m_0 是集合 M 的极限点.完全一样,可以证明立方体是列紧集.直线就不是列紧集,因为如果我们在 ox 轴上取所有点 $1,2,3,\cdots,n,\cdots$ 的集合,这些点以正整数为横坐标,那么这个无穷集将没有任何一个极限点.线段上有理点集也不是列紧集,因为如果我们取一个趋于无理数的有理点序列,那么这个无穷序列在有理数集合中没有极限点.

我们建立列紧集的一些性质如下:

定理 1 (1) 列紧集 R 中的闭集 F 是列紧集.

(2) 所有包含在 R 中的列紧集 F 是 R 中的闭集.(在这里不假定集合 R 本身是列紧集)

证明 (1) 命 M 是集合 F 的任一无穷子集.因为 $F \subset R$,所以 $M \subset R$;因为 R 是列紧集,所以包含在 R 中的无穷集合 M 有极限点 x,此点 x 属于集合 R.点 x 是集合 M 的极限点,更不待说是集合 F 的极限点.但集合 F 是闭集;因此 x 点属于 F.所以 F 中的无穷集 M 有极限点 x,此点 x 也属于 F.这就是说 F 是列紧集.





(2) 现在证明定理的第二个命题. 假设列紧集 $F \subset R$ 不是R中的闭集, 且命 x 是集合 F 的极限点而不属于这集合 F. 在集合 F 中可以找到一个点列

$$x_1, x_2, \cdots, x_n, \cdots$$

此点列以 x 点为其唯一的极限点.

事实上, 以 x_1 表示集合 F 中与x点的距离小于1的点; 以 x_2 表示集合 F 中和 x 点的距离小于下列二数的点

$$\frac{1}{2}, \rho(x, x_1)$$

以 x_3 表示集合 F 中和x点的距离小于下列二数的点

$$\frac{1}{3}, \rho(x, x_2)$$

如此等等.

点 x 将为无穷序列

$$x_1, x_2, x_3, \cdots$$

的极限点, 因为在 x 点的任意邻域中总有此序列的无穷个点存在.

没有其他点 y 是这序列的极限点. 当然, 如果取 n_0 十分大, 使

$$\frac{1}{n_0} < \frac{1}{2}\rho(x, y)$$

则于任意的 $n > n_0$, 对此序列的所有点适合

$$\rho(x, x_n) < \frac{1}{n} < \frac{1}{2}\rho(x, y)$$

由此推出, 对所有的 $n > n_0$

$$\rho(y, x_n) \geqslant \frac{1}{2}\rho(x, y)$$

而这意思是说, 点 y 不是序列

$$x_1, x_2, x_3, \cdots, x_n, \cdots$$

的极限点.

这样一来, 在列紧集 F 中就找到了一个无穷集$x_1, x_2, \cdots, x_n, \cdots$, 此无穷集在这列紧集中没有极限点. 所得到的矛盾说明: 点 x 属于集合F, 即 F 是闭集.

定理2 (1) 列紧集 R 中的闭集的降序列

$$F_1 \supset F_2 \supset \cdots \supset F_n \supset \cdots \quad ①$$

具有非空交集.

(2) 如果集合 F_n 的直径随n增大而趋于零①, 那么集合的交集

$$F_1 \cap F_2 \cap \cdots \cap F_n \cap \cdots$$

① 这意思是说, 无论正数 ε 怎么样, 总可以找到这样的号码 n_0, 使得对于一切 $n > n_0$, $\delta(F_n) < \varepsilon$.

是由一点所构成.

证明 (1) 不限制一般性,我们可以假定序列 ① 的所有集合皆互相不同.由此推出:在集合 F_1 中可以找到一点 x_1 不属于 F_2;在集合 F_2 中可以找到一点 x_2 不属于 F_3;一般在集合 F_n 中可以找到一点 x_n 不属于 F_{n+1}.因为集合 R 是列紧集,那么无穷序列

$$x_1, x_2, \cdots, x_n, \cdots \tag{②}$$

有极限点 x.我们来证明:x 点属于序列 ① 的每个集合中.例如,我们来证明 $x \in F_n$.因为

$$F_n \supset F_{n+1} \supset F_{n+2} \supset \cdots$$

则从 x_n 开始,序列 ② 的所有点皆属于集合 F_n.因为序列

$$x_n, x_{n+1}, x_{n+2}, \cdots \tag{③}$$

和序列 ② 仅差有穷个点,所以序列 ② 的极限点 x 也将是序列 ③ 的极限点.而序列 ③ 的所有点皆包含在 F_n 中,因此 x 点是集合 F_n 的极限,且由于 F_n 是闭集,所以属于 F_n.这样一来,我们就找到了属于序列 ① 的所有集合的点 x.因此这序列的交集是非空集合.

(2) 现在我们假设,集合 F_n 的直径随 n 增大而趋于零.如果序列 ① 的交集包含着两点 x 和 y,那么这两点便属于序列 ① 的每个集合.于是对于所有 n,我们有 $\delta(F_n) > \rho(x, y)$,这就是说,集合的直径不能随 n 增大而趋于零,这与定理的条件相矛盾.因此序列 ① 的集合的交集是由一个点所构成.

点的有穷集合 D 如果具有下面的性质,即对于集合 R 中的每个点 x,总可以找到集合 D 的一点 d,它与 x 点的距离小于 ε,那么集合 D 就叫做集合 R 的 ε - 网.

我们证明:如果 R 是列紧集,那么对于任意 ε,都可以在 R 中找到 ε - 网.假定与此相反,即存在这样的 ε,对这个 ε 列紧集 R 没有 ε - 网.

这意思是说:无论列紧集 R 的一点 a_1 怎样,在 R 中总可以找到一点 a_2,它和 a_1 点的距离大于 ε.因为按照假设,列紧集 R 没有 ε - 网,那么可以找到一点 a_3,它和 a_1 点的距离以及和 a_2 点的距离都大于 ε.因为 a_1, a_2, a_3 诸点不构成列紧集 R 的 ε - 网,那么在 R 中可以找到一点 a_4,它和 a_1, a_2, a_3 每一点的距离都大于 ε.继续做同样的讨论,我们就得到点的无穷序列

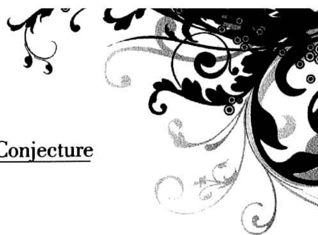
$$a_1, a_2, a_3, \cdots, a_n, \cdots$$

其中每两点的距离大于 ε.但这样的序列是不可能有极限点的.这样一来,我们就在列紧集 R 中找到了一个没有极限点的无穷集 A.这和列紧集的定义相矛盾.因此我们的关于列紧集中对指定的 ε 不存在 ε - 网的假定是不正确的.

所有集合 R,只要具有 ε - 网(对于任一个 ε),则必为有界集.

事实上,命 x 和 y 是集合 R 中的任意两点.在 R 中存在 ε - 网的意思是:可





以找到两点 a 和 b，使

$$\rho(a,x) < \varepsilon, \rho(b,y) < \varepsilon$$

我们用 α 表示集合 R 的 ε - 网的点偶距离之最大者.于是

$$\rho(a,b) \leqslant \alpha$$

因此

$$\rho(x,y) \leqslant \rho(x,a) + \rho(a,b) + \rho(b,y) < \varepsilon + \alpha + \varepsilon = \alpha + 2\varepsilon$$

这就是说集合 R 是有界的.

从证明了的命题立刻推出:列紧集是有界集合.

定理 3　欲使直线上、平面上或空间中的集合 R 是列紧集,其必要而充分的条件是 R 为闭集且为有界集.

必要性　因为集合 R 是列紧集,那么由于定理1的命题(2).R 是一切包含着它的集合中的闭集,所以它是直线、平面或全空间中的闭集,而由于 R 的列紧性,所以 R 为有界集合.

充分性　因为 R 是有界集,所以按照集合 R 是在直线上,或是在平面上,抑或在空间中,其所有点将属于某个线段、正方形或是立方体.集合 R 对于直线、平面或空间为闭集,进而对于包含着它的线段、正方形或立方体亦将为闭集.但是线段、正方形和立方体都是列紧集.因此集合 R 作为列紧集中的闭集,根据定理 1 的命题(1),其本身也是列紧集.

如果集合 R 中的一切开集 G 至少包含集合 D 的一个点,那么集合 D 称为在集合 R 中到处稠密.

因为点的球邻域是开集,所以由此定义立刻可以推出:欲使集合 D 在集合 R 中到处稠密,其必要且充分的条件是:集合 R 的每个点皆是集合 D 的极限点.

例如,所有有理点的集合 D(所有无理点的集合也是一样)在线段上稠密.

定理 4　在一切由无穷个点所构成的列紧集中,有可列的稠密集 D 存在.

欲验明此事,我们考虑集合

$$D = D_1 \cup D_2 \cup \cdots \cup D_n \cup \cdots$$

其中 D_n 是列紧集 R 的 $\frac{1}{n}$ - 网.

无论数 ε 及列紧集 R 的点 x 如何,总有一点 $d \in D$,它和 x 点的距离小于 ε.只要取集合 $D_n\left(此处\frac{1}{n} < \varepsilon\right)$ 中的一点 d 就够了,因此 D 在 R 中稠密.

2.5　连续映象

设 f 是集合 X 到集合 Y 中的一个映象,$x_0 \in X$,如果无论正数 ε 怎样,总可以找到这样的数 δ,使得只要

$$\rho(x,x_0) < \delta$$

那么

$$\rho(f(x),f(x_0)) < \varepsilon$$

于是映象 f 就称为连续映象.

换句话说,如果无论点 $y_0 = f(x_0)$ 的 ε - 邻域怎样,总可以找到 x_0 点的一个 δ - 邻域,使得 δ - 邻域的一切点的象都包含在 y_0 点的 ε - 邻域中,那么映象 f 就在 x_0 点连续.

如果映象 f 在每个点 $x \in X$ 处连续,那么它就叫做集合 X 到集合 Y 中的连续映象.

如果逆映象 f^{-1}(集合 Y 到集合 X 上的映象)也是连续的,那么集合 X 到集合 Y 上的互相单值且连续的映象就叫做拓扑映象.如果二集合中的一个集合可以拓扑映象到另一个集合,则此二集合称为同胚.

我们注意关于连续映象的下列诸定理:

定理 1 欲使集合 X 到集合 Y 中的映象为连续映象,其必要且充分的条件是:每个开集 $H \subset Y$ 的原象为 X 中的开集.

必要性 命映象 f 是连续的,并且集合 $H \subset Y$ 是开集,取集合 H 原象中任一点 x_0.因为 $y_0 = f(x_0)$ 包含在 H 中,那么可以找到 y_0 点的 ε - 邻域也整个包含在 H 中.由于映象的连续性,可以找到这样的数 δ,使 x_0 点的 δ - 邻域的所有点 x 的象将包含在 y_0 点的 ε - 邻域中,因此包含在开集 H 中.因此 x_0 点的 δ - 邻域包含在集合 H 的原象中,所以集合 H 的整个原象是其点的球邻域的和集,因而集合 H 的整个原象是开集.

充分性 事实上,如果条件被满足,那么,点 $y_0 = f(x_0)$ 的 ε - 邻域的原象是包含着 x_0 点的开集 G.所以,可以找到这样的数 δ,使 x_0 点的 δ - 邻域包含在集合 G 中,因此 x_0 点的 δ - 邻域的所有点都映象到 y_0 点的 ε - 邻域中的点,这就证明了映象 f 在 x_0 点的连续性.

因为互余的集合的原象仍互为余集,所以从定理 1 的推出:

定理 1′ 欲使集合 X 到集合 Y 中的映象 f 为连续映象,其必要且充分的条件是:Y 中的每个闭集的原象为 X 中的闭集.

以上我们考虑过的关于集合的许多性质,在连续映象下保持不变.现在我们来证明连通性和列紧性在连续映象之下也是不变的.

定理 2 连通集 X 的连续影象 Y 是连通集.

假设与此相反,即集合 Y 不是连通集,于是它就可以表示为两个互不相交的非空闭集 C 和 D 的和集的形式.根据定理 1′,集合 C 和 D 的原象便是 X 中的闭集 A 和 B.集合 A 和 B 皆非空集,因为 X 被映象到 Y 上.因为集合 C 和 D 无公共点,故 A 和 B 亦不相交.这样一来,我们就把连通集 X 分成两个无公共点的





非空闭集 A 和 B,但这是不可能的.由此推出关于集合 Y 的类似的分割也是不可能的,那就是说,Y 是连通集.

定理 3　列紧集的连续影象是列紧集,设列紧集 X 到集合 Y 的连续映象已被指定.我们应该证明 Y 是列紧集.命 N 是 Y 中的一个无穷集.对于集合 N 的每点 y,我们取这样的一点 x,它是 y 点的原象之一,我们用 M 表示所有这样的点的集合.集合 M 中没有两点被映象于集合 N 的同一点 y.集合 M 和集合 N 同样也是无穷集.因为 X 是列紧集,所以有集合 M 的极限点 x_0 存在.命 $y = f(x_0)$;于是点 y_0 将为集合 N 的极限点.事实上,因为映象 f 在 x_0 点连续,所以无论 ε 怎样,总可以找到这样的 δ,使得如果 $\rho(x, x_0) < \delta$,即 $\rho(y, y_0) < \varepsilon$.因为 x_0 是集合 M 的极限点,所以存在集合 M 中的点 x 的无穷集,这些点和 x_0 的距离小于 δ.于是这些点的影象 y 属于集合 N,且和 y_0 的距离小于 ε.这样一来,对于任意 ε 总可以找到集合 N 中的一点,它和 y_0 的距离小于 ε.这就是说:y_0 是集合 N 的极限点.这样一来,就证明了 Y 的列紧性.

推论　在列紧集 X 到集合 Y 上的连续映象之下,一切闭集(对于 X) 的影象仍为闭集(对于 Y).

事实上,命 A 是列紧集 X 的闭集,B 是集合 A 的影象.按照 2.4 定理 1 的命题(1),集合 A 是列紧集.所以根据刚才证明的定理,集合 B 将为列紧集,而根据定理 1 的命题(2),集合 B 将为 Y 中的闭集,这就是所要证明的事实.

定理 4　局部连通的列紧集的连续影象仍为局部连通列紧集.

证明　命 X 是局部连通列紧集,Y 是集合 X 的连续影象.按定理 3 集合 Y 是列紧集.我们应该证明集合 Y 是局部连通集.

设 y 是集合 Y 的任一点,H 是包含着此点的任意开集.根据定理 1,Y 中开集 H 的全原象 G 是 X 中开集.集合 G 包含着 y 点的全原象 F.

因为集合 X 是局部连通集,那么对于集合 F 的每一点 x,可以找到这样的连通开集 $U(x)$,它包含着 x 而其本身包含在 G 中.我们用 S 表示所有这样的集合 $U(x)$ 的和集,而用 T 表示集合 S 的影响.

而且用 A 表示和集合 S 互余的闭集,用 B 表示 A 的影象.因为 X 为列紧集,而 A 为闭集,所以集合 B 仍为闭集(由于前一定理的推论),所以集合

$$V = Y \setminus B$$

将为 Y 中的开集.

我们证明

$$y \in V \subset T \subset H$$

因为 F 是 y 点的全原象,且

$$F \subset S = X \setminus A$$

所以集合 A 中没有一点以 y 点为其影象,所以

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

$$y \in V = Y \setminus B$$

因为集合 A 的影象是集合 B,而集合 V 是 B 的余集,所以集合 V 中没有一点是集合 A 中某点的影象.但集合 S 是 A 的余集,因此集合 V 的每个点都是集合 S 的某个点的影象.因此集合 V 包含在集合 T 中,T 是集合 S 的影象

$$V \subset T$$

最后因为 $S \subset G$,所以

$$T \subset H$$

这样一来,我们就证明了

$$y \in V \subset T \subset H$$

现在证明集合 T 是连通集.集合 T 可以考虑为所有连通开集 $U(x)$ 的象集的和集,其中 x 是集合 F 中的一点,y 点是此点的影象.于是给出和集 T 的所有的加项都是具有公共点 y 的连通集,所以整个集合 T 是连通集.

用 C_y 表示所有包含着 y 点且被包含于开集 H 中的连通集的和集.

集合 C_y 将是包含着 y 点且被包含于开集 H 中的最大的连通集."最大"这个词的意义应该被了解为什么包含着 C_y 且被包含在 H 中的集合都不是连通集.

我们证明 C_y 是开集.事实上,我们已经看到 y 点属于开集 V,而 V 包含在 T 中,但因为 T 是连通集,所以

$$T \subset C_y$$

这样一来

$$y \in V \subset C_y$$

命 z 是集合 C_y 的任意一点.像上面所说的一样,对于 z 总可以找到连通集 T_z 和开集 W,使

$$z \in W \subset T_z$$

用 C_z 表示包含着 z 点且被包含于集合 H 的最大的连通集,于是我们将有

$$z \in W \subset T_z \subset C_z \subset H$$

集合 C_y 和 C_z 一致.事实上,比如倘若集合 C_z 包含着一个不属于 C_y 的点,那么集合

$$C_y \cup C_z$$

便是包含着 y 点且被包含在集合 H 中而较 C_y 大的一个连通集.但与 C_y 是包含着 y 点且被包含在集合 H 中的最大连通集这样一件事实相矛盾.因此

$$C_y = C_z$$

所以

$$W \subset C_y$$

因为 z 点是集合 C_y 的任意点.那么由此推出,集合 C_y 的所有点都包含在某





个开集中,而此开集本身却又包含在 C_y 中,因此 C_y 可以表示为类似的开集的和集,由此推出 C_y 是开集.

这样一来,对于集合 Y 的任意点 y 和任意包含着 y 点的开集 H,我们找到了连通开集 C_y,适合

$$y \in C_y \subset H$$

这就证明了集合 Y 的局部连通性.

我们已经看到直线线段是局部连通集,而函数

$$\sin \frac{1}{x}, 0 < x < 1$$

图形上的点和极限线段

$$x = 0, -1 \leqslant y \leqslant 1$$

上的点所构成的集合不是局部连通集.

定理 5 列紧集的所有互相单值且在一方连续的映象,在另一方也必连续,所以是拓扑映象.

设 f 是列紧集 X 到集合 Y 上的一个互相单值且在一方连续的映象.由于定理 3,集合 Y 为列紧集.需要证明列紧集 Y 到列紧集 X 的映象 f^{-1},即 f 的逆映象,也是连续的,这就是说,需要证明在映象 f^{-1} 之下,X 中的所有闭集 F 的原象是 Y 中的闭集.然而集合 F 在映象 f^{-1} 之下的原象和集合 F 在映象 f 之下的影象一致,而根据定理 3 的推论,集合 F 在映象 f 之下的影象是 Y 中的闭集.定理就证完了.

我们已经看到,我们所引入的集合的性质中,有许多性质在连续映象之下保持不变.连通性、列紧性、列紧集的局部连通性都在连续映象下保持不变.这些性质在拓扑映象下更加保持不变.但在拓扑映象下还有许多其他的性质(这些性质被一般的连续映象所破坏)仍保持不变.例如,任何集合(不仅是列紧集)的局部连通;集合在包含着它的较大的集合中其为闭集或开集的这种性质;集合成为开集之边界的性质;集合在包含着它的集合中成为稠密集的性质等在拓扑映象下都是保持不变的.

这应该在这种意义下来了解:即如果集合 A 对于包含着它的一个集合 X 而言,它具有某一个所说的性质,而 X 被拓扑映象到集合 Y,那么集合 A 的影象 B 对于 Y 而言,它具有同样的性质.譬如,如果 A 在 X 中为闭集,那么 B 就在 Y 中为闭集;如果 A 在 X 中稠密,那么 B 就在 Y 中稠密.

研究集合在拓扑映象下保持不变的性质,乃是拓扑学的内容.

2.6 连续统的性质

我们今后研究的基本对象将是连续统.同时连通而又列紧的集合叫做连续

统.

上面我们已经证明了线段、正方形、立方体都是连通而列紧的集合,因此它们都是连续统.由2.5的定理2和定理3立刻推出:连续统的连续影象仍为连续统.所以,集合之成为连续统的性质在连续映象下保持不变.这个性质使我们能证明:像圆周、椭圆、双纽线等这样的集合都是连续统,因为其中每个都可以作为有穷个线段所构成折线的连续影象而得到,而由有穷个线段所构成的折线是连续统.一般地说,所有单弧和单闭曲线都是连续统.我们在这一节中建立连续统的某些性质,今后关于曲线的研究就是以这些性质为基础.

定理1 命 C 为连续统序列

$$C_1 \supset C_2 \supset \cdots \supset C_n \supset \cdots$$

的公共部分,序列中的每一项包含在前一项中,那么 C 仍为连续统.

因为集合 $C_1, C_2, \cdots, C_n, \cdots$ 都是列紧集且包含在 C_1 中,所以它们对于 C_1 而言都是闭集.因此它们的交集

$$C = C_1 \cap C_2 \cap \cdots \cap C_n \cap \cdots$$

也是闭集.作为列紧集的降序列的交集,此集合不是空集.所以 C 是列紧集,余下的要证明它是连通集.

集合 C 的连通性的证明基于下列的命题:

如果 C 是列紧集的降序列

$$C_1 \supset C_2 \supset \cdots \supset C_n \supset \cdots \quad ①$$

的公共部分,而 W 是包含着集合 C 的开集,那么在序列①中就有这样的一个列紧集 C_n,它也包含在开集 W 之中.

假定相反,也就是说,没有任何列紧集 C_n 包含在 W 中,于是我们可以构成一个非空闭集序列

$$C_1 \setminus W \supset C_2 \setminus W \supset \cdots \supset C_n \setminus W \supset \cdots$$

其中每一项都包含在前一项中,因为所有这些集合都是列紧集,所以它们的交集不是空集.它是由集合 C 中不属于 W 的那些点所构成.但因为 $C \subset W$,所以这些点是不存在的.由所得到的矛盾,我们应该断定:存在着列紧集 C_n 完全包含在 W 中.

由刚证明的命题我们来道出集合 C 的连通性.假设列紧集不连通.于是它可以表示为两个无公共点的非空闭集 A 和 B 的和集的形式.由于2.2定理5,存在着开集 U 和 V,分别包含集合 A 和 B 且没有公共点.

我们用 W 来表示集合 U 和 V 的和集.因为

$$A \subset U, B \subset V, C = A \cup B, W = U \cup V$$

所以

$$C \subset W$$

根据刚才证明的命题,存在有连续统 C_n 完全包含在 W 中.因为集合 C 既有





点在集合 U 中也有点在集合 V 中,所以集合 C_n 应该既包含有集合 U 的点,也包含有集合 V 的点.这样一来,连续统 C_n 包含在不相交的开集 U 和 V 的和集中,而且同其中每个集合都有公共点.但由于 2.3 定理 1,这是不可能的.这个矛盾就证明了我们假定集合 C 可能不连通是不正确的.因此,C 为连通集,我们的定理也就证明了.

现在我们来证明一个在研究曲线时经常用的定理.这定理本身也是很有趣的.

定理 2 如果 G 是列紧集 R 的连通且局部连通的开集,那么集合 G 的任意两点 a 和 b 总可以用一支完全在集合 G 中的单弧连接之.

证明 我们首先证明:对于任意的正数 ε,集合 G 的两点 a 和 b 总可以用连通开集的一个有穷串链

$$U_1, U_2, \cdots, U_n \qquad ②$$

来连接,这些开集包含在 G 中,其直径小于 ε,而且

$$\begin{cases} a \in U_1, b \in U_n \\ U_i \cap U_{i+1} \neq 0, U_i \cap U_k = 0 \text{ 当 } |i-k| > 1 \end{cases} \qquad ③$$

事实上,由于集合 G 的局部连通性,它的每个点 x 都有连通邻域 $U(x)$,此邻域完全包含在 G 中且其直径小于 ε.我们证明从这许多邻域可以作一个连接 a 和 b 的串链.假定不能这样做.于是我们用 U 表示集合 G 中可以用一串邻域 $U(x)$ 来和 a 点相连接的那些点的全体.

首先我们证明,集合 U 是开集同时也是闭集.因为如果 x 点属于 U,那么 U 就包含着这点的一个邻域,因为此邻域的一切点 y 都可借连接着 a 点和 x 点的那些邻域和邻域 $U(x)$ 所构成的串链与 a 点相连接.所以集合 U 是开集.

因为假若 x 是集合 U 的一个极限点,而 $U(x)$ 是 x 点的连通邻域,那么 $U(x)$ 至少包含集合 U 中的一点 y.点 y 可以用一串链与点 a 相连接.倘若于此串链再添上 x 点的邻域 $U(x)$,就得到一个连接着点 a 和 x 的串链,所以 U 是闭集.

因为集合 U 既是开集也是闭集,所以它的余集

$$V = G \setminus U$$

也既是开集又是闭集.此外因集合 V 包含 b 点,所以不是空集.这样一来,连通集 G 就被分裂为两个无公共点的非空间集(闭集)的和集,而这是不可能的.

由所得到的矛盾,我们应断言:存在着连通开集的有穷序列

$$U'_1, U'_2, , \cdots, U'_5 \qquad ④$$

能使
$$a \in U'_1, b \in U'_s, U'_i \cap U'_{i+1} \neq 0$$

由这序列的元素,我们可以做成适合所有条件 ③ 的一个串链.我们用 U_1 表示序列 ④ 中包含着 a 点且有最大号码的那个元素.用 U_2 表示序列 ④ 中和

U_1 相交而具有最大号码的那个元素．用 U_3 表示序列 ④ 中和 U_2 相交而具有最大号码的元素，循此继进，一直连到第一个包含 b 点的元素 U_n 为止，此处 $n-1$ 是以前所选出的元素的数目．这样一来，我们就获得了一个适合所有要求 ③ 的串链．

应用这个结果来证明我们的定理．为此，我们用一连通开集（其直径小于 1）串链来连接 a 和 b．设这串链是集合

$$U_1, U_2, \cdots, U_n \tag{5}$$

于此

$$\begin{cases} \delta(U_i) < 1 \\ a \in U_1, b \in U_n \\ U_i \cap U_{i+1} \neq 0, U_i \cap U_k = 0, \text{当} \mid i-k \mid > 1 \end{cases} \tag{6}$$

以 c_1 表示集合 U_1 和 U_2 任一交点；c_2 表示集合 U_2 和 U_3 的任一交点；……；c_n 表示集合 U_{n-1} 和 U_n 的任一交点．

现在用直径小于 $\frac{1}{2}$ 的连通开集的串链把 a 点和 c_1 点相连，c_1 点和 c_2 点相连，……，c_{n-1} 点和 b 点相连．在这里我们还要假定连接着 a 点和 c_1 点的串链的构成集合，以及这些集合的闭包，都包含在集合 U_1 中．同样我们假定连接着 c_1 点和 c_2 点的串链的集合，以及这些集合的闭包，都包含在集合 U_2 中，……，最后，连接着 c_{n-1} 和 b 的串链的集合，以及这些集合的闭包都包含在集合 U_n 中．

由这些个别的串链，如果必要，就删去一些多余的元素，我们形成一个连接着 a 点和 b 点而适合条件 ③ 的串链

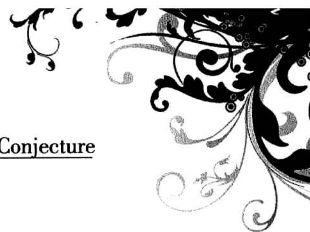
$$V_1, V_2, \cdots, V_p \tag{7}$$

于是，我们用 d_1 表示集合 V_1 和 V_2 的某个交点；d_2 表示集合 V_2 和 V_3 的某个交点；……；最后，用 d_{p-1} 表示集合 V_{p-1} 和 V_p 的某个交点．我们用一个（连通）开集串链将 a 点和 d_1 相连接，这些开集及其闭包皆包含在集合 V_1 中，而其直径小于 $\frac{1}{3}$；再用一个（连通）开集串链来连接 d_1 点和 d_2 点，这些开集及其闭包皆包含在 V_2 中，而其直径小于 $\frac{1}{3}$；……；最后，籍一连通开集串链来连接 d_{p-1} 点和 b 点，这些开集及其闭包皆包含在 V_p 中，而其直径小于 $\frac{1}{3}$．由这些个别的串链，我们又组成了一个连接着 a 点和 b 点的串链．

用同样的办法，我们就得到一系列的连接着 a 点和 b 点的连通开集串链，这系列中第 n 个串链的元素的直径小于 $\frac{1}{n}$，而且每后一串链的诸环节及其闭包皆包含在前一串链的诸环节之中．

我们用 C_1 表示构成第一串链诸集合的和集，用 C_2 表示第二串链诸环节的





和集,如此等等.

集合

$$C_1, C_2, \cdots, C_n, \cdots$$

中的每一个都是连通开集.因此其闭包

$$\overline{C}_1, \overline{C}_2, \cdots, \overline{C}_n, \cdots$$

都是连续统.

因为每后一串链的各环节及其闭包皆包含在前一串链的相应环节中,所以

$$C_{n+1} \subset \overline{C}_{n+1} \subset C_n \subset \overline{C}_n$$

这样一来,集合 $\overline{C}_1, \overline{C}_2, \cdots, \overline{C}_n, \cdots$ 乃构成一连接统的序列

$$\overline{C}_1 \supset \overline{C}_2 \supset \cdots \supset \overline{C}_n \supset \cdots$$

其每后一个连续统皆包含在前一连续统中.根据定理 1,所有这些连续统的交集仍为连续统,我们用 C 来表示它.

因为每个连续统 $\overline{C}_1, \overline{C}_2, \cdots$ 都包含着 a 点和 b 点,所以它们的交集也包含这两点.我们来证明连续统 C 为一连接着 a 点和 b 点的单弧.欲覆验此事,我们应该证明连续统 C 可以互相单值而互相连续地映象到线段

$$T = [0,1]$$

上.

我们把线段 T 分成 n 个相等的线段

$$T_1, T_2, \cdots, T_n$$

称之为一级线段,我们把一级线段对应于第一串链的每个环节.再把每个一级线段分成那么多个相等部分,使得和第一串链中对应于这线段的环节内所包含的第二串链的环节的个数一般多.这样一来,我们得到了 p 个二级线段的串链

$$S_1, S_2, \cdots, S_p$$

再把这些二级线段对应于第二个串链的各环节,把每个二级线段分成那么多个相等部分,就像第二串链中对应于这线段的环节内所包含的第三串链的环节的个数一般多,这就得到了 q 个三级线段的串链

$$R_1, R_2, \cdots, R_q$$

循此继进.

这样一来,我们就得到两个串链序列:第一个序列中的各串链是由列紧集 R 的连通开集所构成;而第二个序列中的各串链是由分线段[0,1]而得到的诸线段所构成.于是,对于每个自然数 n,在第 n 个开集串链和相应的线段串链的诸环节之间有一互相单值的对应存在.

命 x 是连续统 C 的任意点.此点属于第一串链的某个开集 U_k①,也属于第

① 这样的集合可能有两个,但不能更多.

二串链的某个开集 V_l,也属于第三串链的某开集 W_m,如此等等.于是,每后一集合及其闭包皆包含在前一集合中,而且这些开集的半径将随串链的号码增大而趋于零.

但是,第一串链包含着 x 点的环节 U_k,对应着完全确定的线段 T_k;线段 S_l 对应于环节 V_l,而 S_l 包含在 T_k 中;线段 R_m 对应于环节 W_m,而 R_m 包含在 S_l 中,如此等等.这些线段的长度随串链号码增大而趋于零.因此所有这些线段有唯一的公共点 y,而我们就把这一点 y 和连续统 C 的 x 点对应.

连续统 C 的相异二点 x_1 和 x_2 对应于线段 T 的不同的点 y_1 和 y_2.事实上,因为第 n 个串链的环节的直径随号码 n 增大而趋于零,所以可以找这样的号码 N,从这个号码开始,在一切串链中,点 x_1 和 x_2 将属于不同的而无公共点的环节中.于是对应于这两环节的线段亦将没有公共点,因此线段 T 上对应于连续统 C 的 x_1 点和 x_2 点的两点 y_1 和 y_2 也将不相同.

线段 T 的每个点 y 皆对应于连续统 C 的某一点 x.事实上,y 点属于某个一级线段 T_k①,属于某个包含在 T_k 中的二级线段 S_l,属于某个包含在 S_l 中的三级线段 R_m,如此这般.但是线段 T_k 对应着完全确定的第一串链的开集 U_k;线段 S_l 对应着第二串链的开集 V_l,而 V_l 及其闭包皆包含在 U_k 中;线段 R_m 对应着开集 W_m,而 W_m 及其闭包皆包含在 V_l 中,如此等等.这样一来,可以把线段的序列

$$T_k \supset S_l \supset R_m \supset \cdots$$

和闭集的降序列

$$\overline{U}_k \supset \overline{V}_l \supset \overline{W}_m \supset \cdots$$

相对应.于是,线段的长度及其相应的闭集的直径随着串链级数增大而趋于零.

我们用 x 表示属于一切闭集

$$U_k \supset \overline{V}_l \supset \overline{W}_m \supset \cdots$$

的唯一的公共点;线段 T 上所考虑的点 y 就对应于连续统 C 的这一点 x.

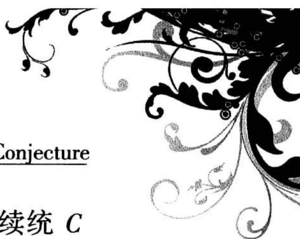
这样一来,我们证明了连续统 C 互相单值地被映象到线段 T 上,其余还要证明这个映象的连续性.

命 x_0 是连续统 C 的任意点,ε 是任意正数.我们用 y_0 表示在所建立的连续统 C 到线段 T 的映象之下,线段 T 上对应于点 x_0 的一点.我们取充分大的 n,使得包含着 y_0 点的 n 级线段②完全界于区间$(y_0 - \varepsilon, y_0 + \varepsilon)$之中,于是,点 x_0 所属的那些开集③的点的总体就被映象到点 y_0 的 ε - 邻域中,这就证明了我们所考虑的连续统 C 到线段 T 的映象的连续性.

由于2.5定理5,因为连续统 C 到线段 T 的连续映象是互相单值的,从而推

① 它可能属于相邻的两个一级线段(是这两线段的公共端点),但不能多于两个.
② 它们可能是一个,也可能是两个.
③ 这样的集合可能是一个或两个.





出线段 T 的连续统 C 上的逆映象的连续性.这样一来,我们就证明了连续统 C 被拓扑映象到区间[0,1]上,此时,点 a 对应于0,而点 b 对应于1,所以连续统 C 是一个单弧 ab.

下面的一个定理有辅助的特性.在许多场合中,运用了这个定理,使曲线理论的说明变得大大地简单了.这个定理可以留到以后用它来证明其他定理时再阅读(见4.3).我们所以在这里把它引进来,是因为它表示出不仅是曲线而是一切连续统所公有的性质.

如果集合 $C\setminus x$ 不是连通集,那么我们就说 x 点分割连续统 C,在相反的情况,即如果集合 $C\setminus x$ 是连通集,我们就说 x 点不分割连续统 C.

定理3 在每个连续统 C 中,至少有两个点不分割连续统 C.

证明 设 a 是连续统 C 的任一点(关于这个点 a,我们暂且不作它是否分割连续统的任何假定).我们证明:有异于点 a 的一点 b 存在,此点 b 分割连续统 C.

设 A 是一个在连续统 C 中稠密而不包含 a 点可列集.

如果集合 A 的任一点 b 不分割连续统 C,那么我们的命题就被证明了.

假定集合 A 的所有点皆分割连续统 C.我们把集合 A 的所有点都标上号码,于是便得序列

$$a_1,a_2,a_3,\cdots \tag{8}$$

按定义,a_1 点将连续统 C 分割为两个(关于 C 的)开集,每个开集以唯一的点 a_1 为其边界点.我们用 A_1 表示这两个开集中不包含 a 点的那一个开集.

其次命 a_{i_2}①是序列⑧中第一个包含在 A_1 中的点.点 a_{i_2} 把连续统 C 分割成两个开集.我们用 A_2 表示其中不包含 a_1 点的那个开集.集合 A_2 的边界由 a_{i_2} 构成.同样,命 a_{i_3} 是序列⑧中第一个包含在 A_2 中的点.它把连续统 C 分割为两个开子集.用 A_3 表示其中不包含 a_{i_2} 点的那个开子集.

作同样的措施,我们就得到开集序列

$$A_1,A_2,A_3,\cdots \tag{9}$$

以及点序列

$$a_1=a_{i_1},a_{i_2},a_{i_3},\cdots \tag{10}$$

每个集合 A_n 的边界由一点 a_{i_n} 所构成.

如果把它唯一的边界点 a_{i_n} 添到集合 A_n 中,那么所得到的

$$\overline{A}=A_n\cup a_{i_n}$$

是连续统.

① $a_{i_1}=a_1$.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

事实上,集合 $\overline{A}_n$ 是闭集,而且作为列紧集 C 的闭集,它是列紧集.余下的是要证明集合 $\overline{A}_n$ 是连通集.假定相反,设 P 和 Q 是两个无公共点的非空闭集,其和集等于 $\overline{A}_n$.假设 a_{i_n} 点属于集合 Q,于是

$$P \subset A_n$$

因为集合 $C \setminus A_n$ 和 Q 每个都是闭集,所以其和集 $(C \setminus A_n) \cup Q$ 也是闭集.这样一来,我们就把连续统 C 分割为两个无公共点的非空闭集 P 和 $(C \setminus A_n) \cup Q$ 的和集,而这是不可能的.所以我们应该断言:集合 $\overline{A}_n$ 是连通集.

集合 $\overline{A}_n$ 既是列紧的而且也是连通的,所以是连续统.我们考虑连续统的序列

$$\overline{A}_1, \overline{A}_2, \cdots, \overline{A}_n, \cdots$$

并且证明每后一连续统皆包含在前一连续统中

$$\overline{A}_{n+1} \subset \overline{A}_n$$

事实上,按照假定,点 a_{i_n} 把连续统 C 分成两个开集 A_n 和 B_n.因为集合 $\overline{A}_{n+1}$ 不包含 a_{i_n} 点,所以它包含在集合 A_n 和 B_n 的和集中.但因为 $\overline{A}_{n+1}$ 是连通集,所以它只能和集合 A_n 和 B_n 中的一个集合有公共点;在 $a_{i_{n+1}} \in A_n$,故 $\overline{A}_{n+1} \subset A_n \subset \overline{A}_n$.

这样一来,我们就得到连续统的序列

$$\overline{A}_1 \supset \overline{A}_2 \supset \cdots \supset \overline{A}_n \supset \cdots$$

其中每后一连续统皆含于前一连续统中.根据定理 1,所有这些连续统的公共部分仍是连续统,我们用 K 来表示它.由于闭包

$$\overline{A}_{n+1} \subset A_n$$

连续统 K 可以看做开集 $A_1, A_2, \cdots, A_n, \cdots$ 的交集.

设 b 是连续统 K 的任一点.因为没有一点 $a_{i_m}(m = 1,2,\cdots,n)$ 是属于连续统 $\overline{A}_{n+1}$ 的,由此推知,b 点不和任何一点 a_{i_n} 相一致.

我们来证明 b 点不分割连续统 C.假定相反,即 $C \setminus b$ 是两个无公共点的开集 G 和 H 的和集.

因为 b 点不与序列 ⑩ 的任一点相一致,所以这序列的每一点属于集合 G 或 H 中的一个集合.譬如,我们假定序列 ⑩ 的某一点 a_{i_m} 属于 H,我们证明在这时集合 $\overline{G} = G \cup b$ 完全包含在序列 ⑨ 中对应于 a_{i_m} 点的集合 A_m 内.首先注意,集合

$$\overline{G} = G \cup b$$

是连续统.与此完全一样,可以证明每个集合

$$\overline{A}_n = A_n \cup \alpha_{i_n}$$





都是连续统.

集合 $\overline{G}$ 不包含 a_{i_m} 点,并且与集合 A_m 有公共点 b. 所以 $\overline{G}$ 完全包含在集合 A_m 中,因为它是连通的,所以它不可能和两个互不相交的开集 A_m 和 B_m 有公共点.

现在我们证明集合 $\overline{G}$ 完全包含在序列 ⑨ 的每一集合 A_n 中,为此必须证明:如果只要序列 ⑩ 的一个点 a_{i_m} 属于集合 H,那么序列 ⑩ 的所有点都属于这个集合.

与此断言相反,我们假定序列 ⑩ 的任一点 a_{i_p} 属于集合 G. 于是,和以前一样,我们证明集合

$$\overline{H}=H\cup b$$

完全包含在序列 ⑨ 中相应于 a_{i_p} 点的集合 A_p 内.

不限制一般性,可以假设 $p>m$. 因为根据已证明了的事实

$$A_p\subset A_m$$

所以由关系式

$$\overline{H}\subset A_p$$

可以推出

$$\overline{H}\subset A_m$$

这样一来,我们有

$$\overline{G}\subset A_m,\overline{H}\subset A_m$$

注意到

$$\overline{G}\cup\overline{H}=C$$

由此得到

$$C\subset A_m$$

这是错误的,因为集合 A_m 不包含(例如)a 点.

由这个矛盾,我们应断定:如果序列 ⑩ 有任何一点 a_{i_m} 属于集合 H,那么序列 ⑩ 的所有点都属于这个集合. 但是,正如我们以上证明的,由此推出集合 $\overline{G}$ 包含在序列 ⑨ 的每个集合 A_n 中,那么 G 就更加属于序列 ⑨ 的每个集合 A_n 中. 在这种情况下,集合 G 包含在连续统 K 中,K 是序列 ⑨ 的一切集合 A_n 的交集.

由开集 G 是非空集合这样一个假定推出:序列 ⑧ 有一点 a_k 属于集合 G,因为由集合 A 的元素我们构成了一个在 C 中到处稠密的序列 ⑧. 我们在序列 ⑩ 中,取这样的一点 a_{i_n},这一点的号码 i_n 大于点 a_k 的号码 k. 我们便有

$$a_k\in G\subset\overline{G}\subset K\subset A_{n-1}$$

就是

$$a_k\in A_{n-1}$$

但是按照假设,a_{i_n} 是序列 ⑧ 中包含在 A_{n-1} 内而具有最小号码的一点,但我们

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

找到了序列⑧的一点 a_k 属于A_{n-1}，而且它的号码 $k < i_n$. 由这一个矛盾，我们不得不作结论说，b 点不分割连续统 C.

这样一来，在连续统 C 中我们找到了一点 b，它不分割 C. 欲证明存在另一点也不分割连续统 C. 应逐字地重复上面的推论，仅仅不是从任意点 a 出发进行论证，而是在整个推论中用 b 点来代替a 点，关于 b 点我们已经知道它不分割连续统 C. 这就证明了定理 3.

线段恰有两点不分割它的连续统的实例. 线段的端点是不分割的点；而线段的一切内点皆把线段分割为两个以此点为公共端点的半区间.

圆周可以作为其任一点也不分割它的连续统的实例.

3　康托曲线

在这一部分中，我们考虑平面上的集合；在这里，平面起着集合 R 的作用. 所以，读到开集和闭集，我们是指它们为平面上的闭集或开集而言.

康托把(平面)曲线定义为这样的连续统，在这个连续统的每点的任意邻域里，有不属于这个连续统的平面的点存在. 换句话说，具有下面的性质的连续统 C 叫做曲线：无论连续统 C 的一点 x 和正数 ε 如何，在平面上可以找到一点 y，它不属于连续统 C 而且与 x 点的距离小于 ε，从曲线的这个定义直接推出：连续统 C 为康托曲线的充分且必要条件是：它不包含任何开子集. 直到现在我们所考虑过的那些平面曲线(直线、圆周等)都适合这个定义. 反之，如三角形、圆、正方形(如果考虑到它们的内点及边界点)等连续统就不是康托定义的意义下的曲线，因为其每一个都包含一个充分小的球邻域(这里的球邻域是指圆的内点的总体).

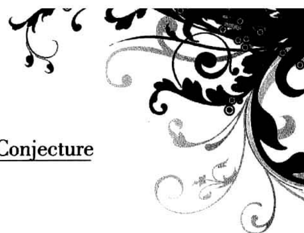
我们再引进康托曲线的一个精彩的例子，这个例子是波兰数学家希尔宾斯基(Sierpinski)所作，以后就叫它为“希尔宾斯基覆盖”(Ковра Серпинского).

将正方形 Q 分为 9 个相等的正方形，并挖去它们中心正方形的内点. 再把剩下的 8 个一级正方形的每一个分为 9 个相等的正方形，并且挖去它们当中正方形的内点. 这样一来，我们就得到 $8^2 = 64$(个)二级正方形. 对于其中每一个作如上之处理，使得 $8^3 = 512$(个)三级正方形，循此继进以至对任意自然数 n(图 19.8).

施行了这一切手续以后，剩下的集合 S' 便是希尔宾斯基覆盖.

在康托定义的意义下，这集合是曲线. 首先因为集合 S' 是一系列连续统的公共部分：其第一个连续统是由 8 个一级正方形所构成，第二个是由 64 个二级正方形所构成，……，此外，在这序列中，每后一个连续统都包含在前一个连续





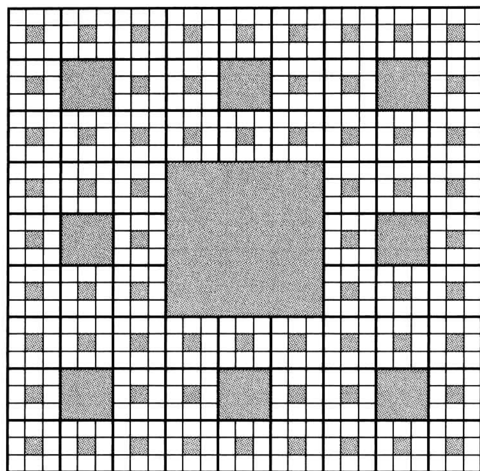


图 19.8

统中①,所以集合 S' 是连续统.

其次,在连续统 S' 的每个点的任意邻域中,都可以找到一点不属于这个连续统.

事实上,无论连续统 S' 的一点 x 以及正数 ε 如何,总可以找到自然数 n,使得 n 级正方形 Q^n 的对角线小于 ε,于是就有相邻于 Q^n 的被挖掉的正方形的点落在 x 点的 ε - 邻域里,这些点不属于连续统 S'.

因此,希尔宾斯基覆盖确实是康托曲线,这曲线具有一个美妙的性质:就是它包含任意的平面曲线.也就是说,下面的定理成立.

定理 无论康托曲线 C 如何,在希尔宾斯基覆盖中总可以找到一个子集 C',这个子集 C' 和集合 C 同胚.

证明 因为集合 C 是连续统,故为有界闭集,所以有包含着连续统 C 的长方形 R 存在,我们把长方形 R 用平行于它的边的直线分成 9 个相等的长方形(在图 19.9 中,这些直线是用虚线给出的).因为 C 是康托曲线,所以它不包含任何开子集,特别它就不能以整个当中的长方形为其子集.因为 C 是闭集,所以在当中的长方形内部包含着一个长方形 R_0,它的边平行于长方形 R 的边,而且它不包含曲线 C 的任何一点.延长长方形 R_0 的各边一直到和基本长方形 R 的边相交为止,于是整个长方形 R 被分为 9 个长方形.我们从长方形 R 中挖掉长方形 R_0 的所有内点.剩下的集合是由 8 个长方形所构成,我们称这 8 个长方形为一级长方形,我们按照逆时针围绕的顺序,从左边最下面的一个长方形开

① 参考 2.6 定理 1.

始,分别用 $R_1^1, R_2^1, R_3^1, \cdots, R_8^1$ 来表示这八个长方形.我们用 S^1 来表示属于诸一级长方形的点的总体.

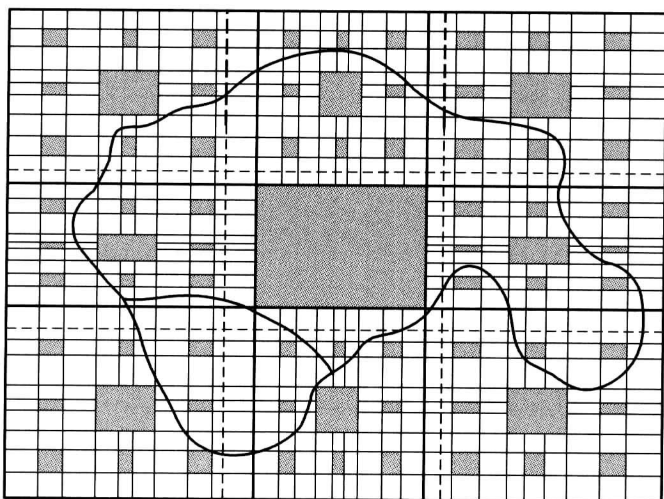


图 19.9

欲得二级长方形,我们处理如下:将每个一级长方形分为 9 个相等的长方形.我们取长方形 R_1^1,于是按照前面的办法,在分割长方形 R_1^1 时获得的当中的长方形的内部,可以找到这样的长方形 P_{10}^1,此长方形的内部不包含曲线 C 的任何一点.延长它的各边一直到和基本长方形 R 的各边相交为止.用同样的办法,在分割长方形 R_2^1 时获得的当中长方形的内部,也可以找到这样的长方形 P_{20}^1,此长方形位于刚才建立的长方形 P_{10}^1 边的延长线之间带域中,而且它不包含曲线 C 的任何点.我们再延长长方形 P_{20}^1 的各边一直到它们和基本长方形的诸边相交为止.完全一样,在长方形 R_3^1 是由分割 R_3^1 时所形成的中心长方形的点所构成,不过我们要求这些点都包含在延长长方形 P_{10}^1 和 P_{20}^1 的诸边而得到的带域的公共部分内.同样,我们得到长方形 $P_{40}^1, P_{50}^1, P_{60}^1, P_{70}^1, P_{80}^1$.

我们用 R_{k0}^1 表示包含在 P_{k0}^1 中而且由所有纵带域和横带域的交集的内点所构成的长方形,这些带域是由延长长方形 P 的边而形成并且穿入 R_k^1 中,让我们来考虑藉延长长方形 R_{k0}^1 的边而形成的基本长方形的一个分割.

如果从每个一级长方形 R_k^1 中挖掉它属于长方形 R_{k0}^1 的内点,那么由每个一级长方形就剩下 8 个长方形,我们称它们为二级长方形,并且按逆时针围绕顺序仍由左下方的长方形开始用 $R_{k1}^2, R_{k2}^2, \cdots, R_{k8}^2$ 来表示.属于所有 64 个二级长方形的点的总体用 S_2 来表示.

像以前我们对一级长方形的措施一样,我们对每个二级长方形 $R_{kl}^2(k, l = 1, 2, \cdots, 8)$ 作同样的措施.这样一来,我们就得到 $8^3 = 512$(个)三级长方形



$R^3_{klm}(k,l,m=1,2,\cdots,8)$，我们用 S_3 来表示这些三级长方形的点的总体.

进而，作同样的处理，我们就得到一连续统序列 $S_1\supset S_2\supset\cdots\supset S_n\supset\cdots$，其每后一个包含在前一个中，根据2.6定理1，所有这些连续统的交集仍为连续统，我们用 S 来表示它. 曲线 C 完全包含在连续统 S 中. 我们来证明连续统 S 和希尔宾斯基覆盖同胚. 命 x 是连续统 S 的任意一点. 那么按连续统 S 的定义，点 x 属于集合 S_1，因而属于某个一级长方形 R^1_k；同时点 x 属于集合 S_2，因而属于某个包含在 R^1_k 中的二级连续统 R^2_{kl} 中；同时点 x 属于集合 S_3，因而属于某个包含在 R^2_{kl} 中的三级长方形 R^3_{klm} 中. 这样一来，我们就得到长方形的降序列 $R^1_k\supset R^2_{kl}\supset R^3_{klm}\supset\cdots$

把构成希尔宾斯基覆盖的一级正方形 Q^1_k 对应于长方形 R^1_k，这正方形 Q^1_k 在基本正方形 Q 中所占的位置和长方形 R^1_k 在长方形 R 中所占的地位一样(参看图19.8和图19.9)，其次把二级正方形 $Q^2_{kl}\subset Q^1_k$ 对应于二级长方形 R^2_{kl}，这个正方形 Q^2_{kl} 在 Q^1_k 中所占的地位和长方形 R^2_{kl} 在长方形 R^1_k 中所占的地位一样，如此继续下去.

这样一来，我们就把正方形的降序列

$$Q^1\supset Q^2\supset\cdots\supset Q^n\supset\cdots$$

对应于长方形的降序列

$$R^1\supset R^2\supset\cdots\supset R^n\supset\cdots$$

但因为 n 级正方形的边长等于 $1/3^n$，所以(根据2.4定理2) 正方形的交集

$$Q^1\cap Q^2\cap\cdots\cap Q^n\cap\cdots$$

不是空集合而且是由一点所构成，而用 x' 来表示这一点. 我们就把这一点 x' 对应于连续统 S 的 x 点.

这样一来我们便证明了，连续统 S 的每一点 x，对应着希尔宾斯基覆盖的某一点 x'. 现在我们证明：连续统 S 的不同的点，对应着连续统 S' 的相异之点. 命 x 和 y 是连续统 S 的两点. 因为每后一级的长方形的边总比前一级包含着它的长方形的相应边小2/3倍，所以随级数的增大，长方形的边长趋于零，因而其对角线亦趋于零. 由此推出，无论连续统 S 的两点如何，总可以找到充分大的号码 n，使得包含着 x 点和 y 点的两个 n 级长方形没有公共点. 于是相应于这两个长方形的正方形也将没有公共点. 因此，对应于 x 点和 y 点的 x' 点和 y' 点也不能互相重叠.

我们已经证明了：连续统 S 的每一点，对应着连续统 S' 的某一点，并且连续统 S 的不同的点对应着连续统 S' 相异之点. 完全一样，也可以证明：连续统 S' 的每一点，对应着连续统 S 的一个(而且只有一个) 点. 因此，在连续统 S 和 S' 之间所建立的对应是互相单值的. 剩下还要证明这对应是互相连续的.

命 x_0 是连续统 S 的任一点，ε 是任一正数，在上面建立的连续统 S 到连续统 S' 的映象之下，x_0 是点 x 的影象. 必须证明有这样的数 δ 存在，使得如果

$$\rho(x, x_0) < \delta$$

那么

$$\rho(x', x'_0) < \varepsilon$$

其中 x' 是点 x 的影象.

我们把 n 取得如此大，使得包含 x'_0 点的 n 级正方形完全界于此点的 ε - 邻域中，我们取一个中心在 x_0 点的圆，它的半径为 δ，此圆完全在包含着 x_0 点的诸 n 级的长方形的和集合，于是当

$$\rho(x_0, x) < \delta$$

即

$$\rho(x'_0, x') < \varepsilon$$

同样可以证明连续统 S' 到连续统 S 的映象的连续性，因此连续统 S 和 S' 是互相同胚的.

但是曲线 C 完全包含在边续统 S 中. 在连续统 S 到希尔宾斯基覆盖 S' 的同胚映象之下，曲线 C 同胚地映象到某一个包含在 S' 中的集合 C'. 这样一来，就证明了我们的定理.

按语 （平面上的）集合成为康托曲线的性质，在拓扑映象下保持不变. 这就是说：如果康托曲线 C 被拓扑映象到平面点集 C'，那么集合 C' 也是康托曲线. 集合的连通性和列紧性不仅在拓扑映象下保持不变，甚至在连续映象下也保持不变①，由此推出：集合 C' 是连续统. 证明中最困难的一部分是下面断言的证明：即如果平面点集 C 不包含任何开子集，那么和这个集合 C 同胚的平面点集 C' 也不包含任何开子集. 这可以从下面的定理推出，如果 G 是平面上的开集，H 是与集合 G 同胚的平面点集，那么 H 也是平面上的开集. 在这里我们不来证明这个定理②.

并不是所有康托曲线都可以用参数方程

$$x = \varphi(t), y = \psi(t)$$

给出，以上再三提到的由函数

$$y = \sin\frac{1}{x}, \quad 0 < x \leqslant 1$$

的图形和极限线段

$$x = 0, \quad -1 \leqslant y \leqslant 1$$

的点所构成的集合，便是这种不能用参数方程给出的曲线的实例. 这个连接统不包含任意开子集，所以是康托曲线. 在 2.5 已经证明了这个曲线不能表示为

① 参看 2.5，定理 2 和定理 3.

② 对于任意维空间的情形，本定理的证明见 П · С · 亚历山大洛夫著的组合拓扑学（Комбинатория тoпoдогия），Гостехиздат，М.—Л. 1947，第 196 页.





线段的连续影象.由此推出:它是不能用参数方程给出的.这种不能用参数方程来表示的原因是在于:所考虑的这曲线不是局部连通集.

可以证明:欲使一个连续统,特别是康托曲线为一线段的连续影象,充分且必要的条件是:它为局部连通集.

此条件的必要性直接由2.5定理4推出;因为线段是局部连通的连接统,而在连续映象之下,一个连续统为局部连通这一性质是保持不变的.

在这里我们不证明这个条件的充分性.因为这个证明可在 Хаусдорф 的 *Теория множеств* 第 186 页找到①.

局部连通的性质可以用另外一种方式表达出来,我们提一提(但不证明)下面这个定理:欲使连续统 C 是局部连通集,充分且必要的条件是:对于任意的 ε,它可以表示为有穷个直径小于 ε 的连续统的和集的形式.

特别,希尔宾斯基覆盖具有这种性质.如果我们取充分高阶的正方形,我们就可以把希尔宾斯基覆盖分割成任意小的连续统.希尔宾斯基覆盖界于每个这样的正方形中的部分也是希尔宾斯基覆盖,不过是在这个正方形上建立起来的.我们把希尔宾斯基覆盖表示成这些部分(其每个都是连续统)的和集的形式.由这种表示的可能性推出:希尔宾斯基覆盖是局部连通的集合,因此它可以表示为线段的连续影象,所以可以用参数方程来给出.这样一来,希尔宾斯基覆盖不仅是康托意义上的曲线,而且是若当意义上的曲线.

由于康托曲线的定义极其广泛,因此,它们可能有很复杂的结构.我们将联系着平面区域的边界来考虑这个问题.

如我们在初等几何学中已经知道的,圆周把平面分为内外两个区域,并且它是这两个区域的公共边界.

不仅是圆周,而且是所有和圆周同胚的曲线,都具有把平面分割为两个区域的特性,这就是著名的若当定理的内容.这个定理,不要看它表面上很明确,但是证明起来却够难的呢!贝努利的双叶线(Лемниската Бернулли)就把平面分成三个区域,而且这曲线本身的交点 s 属于这三个区域的边界(图 19.10).

然而这曲线的其余各点,则仅是某两个区域的边界.骤然看来,似乎曲线上只有个别的点才能属于两个以上的区域的边界.事实上,在地图上,有某一曲线一下子就定出三个或更多个国家的疆界,虽然这很难想象,但这样的曲线确实是存在的,也就是离这种曲线的每一点不管多么近,总有三个或更多个区域的点存在,在这种意义下,这曲线就叫做这些区域的交界.

为了了解这怎么可能,我们想象在大海中有一个岛,在这岛上有两个湖,且想象下面的施工程序是:在第一小时中,从海和每个湖开凿河流,使每条河流都

① Hausdorf 著的 Mengenlere 的俄译本.—— 原译者

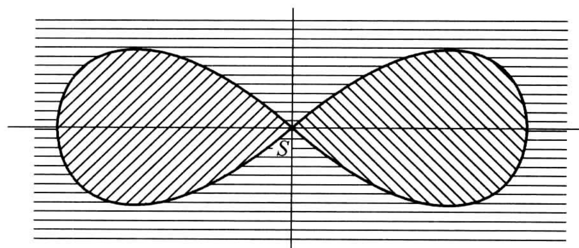


图 19.10

是“盲” 的(那就是说:这些河流实际上都是其对应水泊的湾),使这些河流决不相汇合,而且使得,由于这一小时的施工而从陆地的每一点到达海水以及到达每个湖水的距离小于 1 公里.在其后的半小时中,再把这三条已经开凿的河流继续延长,使得所有河流仍然是“盲” 的,而且不相适合,并使从陆地的每一点到这三个水泊的水的距离小于 1/2 公里.在以后的 1/4 小时中,继续延长这些河流,使它们仍然不相连接,并使它们渗入该岛内部如此地“稠密”,以至从陆地每一点到这三水泊的水的距离都变得小于 1/8 公里,循此继进.这样施工了两小时以后,这个岛仅剩下一个这样的连续统 C,无论离这连续统 C 的每点多么近,总可找到这三水泊的水,而且这三水泊(海和两个湖) 仍然是分开的,即其中任意二水泊的水都不相渗混.这三个(藉引自它们河流而延伸了的) 水泊乃是这样的三个区域,连续统 C 构成其公共交界;其中有一个区域(“海”) 是无界的,其余的两个都是有界的(图 19.11)①.

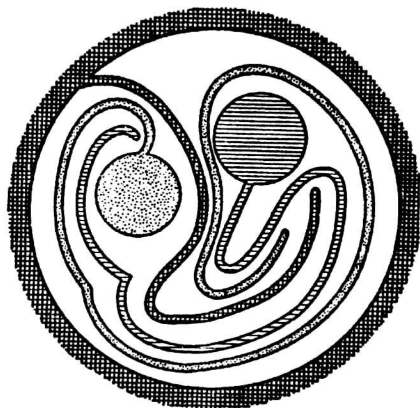


图 19.11

我们这样仔细地读康托曲线,是因为如果仅限于平面曲线,则在我们的数学实际中,牵涉曲线概念者,康托的定义都完全一般地表达了出来.但我们想给出更一般的曲线定义,这定义不仅概括平面曲线,而且也概括空间曲线②.

不可能把康托曲线定义简单地搬到空间上来,因为,如正方形在空间便是一个不包含任何开集的连续统.这就是说,我们又需要来回答“什么是曲线” 这个

① 这个例是与归之于荷兰数学家 Brauwer.这一说明是从 Π · C · 亚历山大洛夫著的《集与函数的泛论初阶》中摘录下来的.

② 甚至是维度更高的空间中的曲线.





问题了.现在我们转到乌利松所给出的曲线概念的定义上来.

4 曲线的一般定义

4.1 曲线的定义及基本性质

我们来看一看,我们管它叫曲线的那些集合具有些什么样的一般的性质?

首先,这些集合都是连通集,不十分严格地说,即它们是由一整块所构成,而不能分割为孤立的部分.

其次,我们所考虑的很多曲线都是有界闭集,所以是列紧集①.直线线段、圆周、椭圆、贝努利双曲线等都是这样的曲线.

今后我们只把那些列紧的连通集,即连续统考虑为曲线②.

但是像正方形、立方体这样的集合也是连续统.为什么我们不叫它们为曲线呢?因为它们具有二维度或三维度.曲线只是具有"一维度"的连续统,但是"具有一维度"这句话是什么意思呢?在这里我们不来讨论什么叫做维数的问题,或如上所说,任意集合的维度③的问题,我们仅仅给出一维连续统的严格定义.

如果有 x 点的任意小邻域存在,此邻域的边界不包含任何由多于一点所构成的连续统,那么我们说,连续统 C 在点 x 具有维度 1.换句话说,如果对于所有 $\varepsilon > 0$,存在着包含 x 点的开集,其直径小于 ε,而其边界不包含任何由多于一点所构成的连续统,那么连续统 C 在 x 点具有维度 1.

如果连续统 C 在它的每一点都是一维的,那么我们就说,连续统 C 具有维度 1.

现在我们可以说:一维的连续统叫做曲线,这就是说,那样的连续统,其每一点都有任意小的邻域,此邻域之边界不包含任何由多余一点所构成的连续统,这样的连续统叫做曲线,这就是乌利松所给出的曲线定义.

线段是一维连续统最简单的实例,也就是乌利松意义上的曲线的实例.事实上,无论线段的任意一点的 ε - 邻域如何,此邻域总是一个开集,如果 x 是线段的内点,则此开集的边界是由两点所构成,如果 x 是线段的端点,则此开集的边界是由一点所构成.

---

① 参看 2.4,定理 3.

② 诚然,许多几何对象也叫做曲线,但是不具有上述的列紧性和连通性.首先全直线便是如此,它不具有列紧性.其次,双曲线就既不是列紧集,也不是连通集.但是我们所以在这里只考虑其为连续统的那些曲线,是因为这种情况最重要,而且在数学实践中叫做曲线的那些几何对象,即使不是连续统,实质上也可以化为上述的情形来考虑.

③ 这个字在俄文中是 измерение 或 размерность,在这里译作"维"、"度"或"维度".通常所讲的"一度空间"、"二度空间"等术语中的"度"就是这里所讲的.—— 译者

同样可以相信,圆周在它的每一点具有维度 1,那就是说,在乌利松的意义下,圆周是曲线.只要仍旧取圆周任意点 x 的 ε - 邻域就够了.此邻域都是以 x 点为中心(不包含端点)的弧.它是开集,其边界是由两点(弧的端点)所构成.

一般地,所有康托曲线都是乌利松曲线.

在证明这个命题之前,我们先注意两件事情,这两个事实不仅在证明上面的命题时是有用的,并且在其他的许多情况,也很有用.

定理 1 如果连续统 K 是(乌利松意义下)曲线 C 的子集,那么它本身也是曲线.

事实上,设 x 是连续统 K 的任意点,ε 是任意正数.因为 C 是曲线,所以有开(对于 C)集合 G 存在,此开集 G 包含 x 点而其直径小于 ε,并且其边界不包含任何连续统.我们用 U 来表示开集 G 中属于连续统 K 的那些点的总体.集合 U 是(关于 K 的)开集.它包含 x 点,其直径小于 ε,又因为其边界(对于 K 取的)是开集 G 的边界的子集,故不包含任何连续统,所以 K 是曲线.

定理 2 曲线的拓扑影象仍然是曲线.

事实上,设 X 是一维的连续统,f 是集合 X 到集合 Y 上的拓扑映象,我们应该证明 Y 仍然是一维的连续统.

因为连通性以及列紧性的一切连续映象之下都保持不变,当然在拓扑映象之下也保持不变,由此推出 Y 是连续统.剩下还要证明这连续统是一维的.

命 y_0 是连续统 Y 的任意一点,ε 是任一正数.我们用 x_0 表示 y_0 点的原象,而且 δ 表示这样的正数,使得如果

$$\rho(x_0, x) < \delta \quad 且 \quad y = f(x)$$

那么

$$\rho(y_0, y) < \varepsilon$$

因为 X 是曲线,所以可以找到包含着 x_0 点的这样的开集 U,它本身包含在此点的 δ - 邻域中,而且集合 U 的边界不包含任何由多于一点所构成的连续统.我们用 V 表示集合 U 的影象.集合 V 是开集(对 Y 而言),它包含 y_0 点且包含在 y_0 的 ε - 邻域中.

我们证明集合 V 的边界不包含任何由多于一点所构成的连续统.因为开集 V 是闭集 U 在拓扑映象 f 下的影象,所以集合 V 的边界是集合 U 的边界的影象.倘若集合 V 的边界包含某个由多于一点所构成的连续统 K,那么在连续映象 f^{-1} 之下(f^{-1} 是映象 f 的逆映象),此连续统的影象是连续统 C,C 也是由多于一点所构成而且包含在开集 U 的边界中,这与集合 U 的定义相矛盾.由此我们应该断定开集 V 的边界不包含任何由多于一点所构成的连续统.

这样一来,我们就证明了连续统 Y 的任一点 y_0 都具有任意小的邻域 V,这邻域的边界不包含任何由多于一点所构成的连续统.这就证明了连续统 Y 是曲线.





预先作了这些考察以后,我们来证明上面所说的关于所有康托曲线都是乌利松曲线的断言.

事实上,所有康托曲线 C 都和希尔宾斯基覆盖 S 的一个子集 C' 同胚.因为 C 是连续统,所以集合 C' 也是连续统.如果我们证明希尔宾斯基覆盖 S 是乌利松意义上的曲线,那么根据定理1,由此推出连续统 C' 是乌利松意义下的曲线.于是由于定理2,连续统 C 是(乌利松意义上的)曲线,所以我们需要证明:

希尔宾斯基覆盖是曲线(乌利松意义上).

在前面,我们已经看到希尔宾斯基覆盖 S 是连续统.我们需要证明:无论连续统 S 的一点 m_0 如何,总存在着包含 m_0 点的任意小的开集,而其边界不包含任何由多于一点所构成的连续统.

我们用 x_0 和 y_0 表示 m_0 点的坐标系中的直角坐标,此坐标系的轴是沿着正方形 Q 的边,我们的希尔宾斯基覆盖 S 就是从 Q 得来的.我们设正方形的边长等于1.

命 ε 是任意正数,α 是适合下列条件的一个正数:$\alpha < \frac{\varepsilon}{2}$,而且四数 $x_0 - \alpha$,$x_0 + \alpha$,$y_0 - \alpha$,$y_0 + \alpha$ 中没有一个能分解成有穷的三进小数①.四直线

$$x = x_0 - \alpha, x = x_0 + \alpha$$

$$y = y_0 - \alpha, y = y_0 + \alpha$$

构成一个正方形,此正方形的内点的总体是由下列不等式所确定

$$x_0 - \alpha < x < x_0 + \alpha, y_0 - \alpha < y < y_0 + \alpha$$

因为 $\alpha < \frac{\varepsilon}{2}$,所以

$$x_0 - \frac{\varepsilon}{2} < x_0 - \alpha < x_0 < x_0 + \alpha < x_0 + \frac{\varepsilon}{2}$$

$$y_0 - \frac{\varepsilon}{2} < y_0 - \alpha < y_0 < y_0 + \alpha < y_0 + \frac{\varepsilon}{2}$$

由此推出点 $m_0(x_0, y_0)$ 是正方形的内点.整个正方形是在以 m_0 点为圆心以 ε 为半径的圆内.每个直线

$$x = x_0 - \alpha, x = x_0 + \alpha$$

$$y = y_0 - \alpha, y = y_0 + \alpha$$

与希尔宾斯基覆盖相交于康托完全集.

---

① 我们说一个数 p,$0 \leqslant p \leqslant 1$,可以分解成有穷的三进位小数,意思是:如果它能表示为

$$p = \frac{p_1}{3} + \frac{p_2}{3^2} + \cdots + \frac{p_n}{3^n}$$

其中诸数 $p_1, p_2, \cdots, p_n$ 可以取数值0,1,2.例如数

$$\frac{7}{27} = \frac{0}{3} + \frac{2}{3^2} + \frac{1}{3^2}$$

便是如此.

如果取我们所考虑的正方形中属于希尔宾斯基覆盖 S 的那些内点的总体，那么便得到 S 中的一个开集，它包含 m_0 点而且被包含在此点的 ε – 邻域中，而其边界不包含任何由多于一点所构成的连续统. 这就证明了希尔宾斯基覆盖是乌利松意义上的曲线. 同时也证明了所有康托曲线都是乌利松意义上的曲线.

逆定理也是正确的，那就是说：所有乌利松意义上的平面曲线都是康托曲线. 这个命题的证明归之于：没有任何乌利松意义上的曲线是包含开集(对于平面)的.

如果曲线 C 包含了平面上的一个开集，那么它就包含一个完全界于这开集中的圆. 此圆为连续统(曲线 C 的子集)，所以根据定理 1，它便是乌利松意义上的曲线.

因此，我们的任务便是要证明任何圆(其内点和边界点都计算在内)都不是乌利松意义上的曲线.

假定相反，命 K 是某个圆的内点的总体，x 是圆的任一内点，G 是包含着 x 点的一个开集，此开集及其闭包都是由圆的内点所构成，而且其边界

$$D = \Gamma p(G) = \overline{G} \setminus G$$

不包含任何由多于一点所构成的连续统.

但我们已看到：如果 G 是某集合 R 的非空间子集，而和 R 不相一致，那么这开集的边界 $\Gamma p(G)$ 就分割整个集合 R，这意义是说：如果从集合 R 中去掉开集 G 的边界，那么剩下的集合

$$R \setminus \Gamma p(G)$$

表示为两个开集 G 和 $R \setminus \overline{G}$ 的和集的形式，而这两开集 G 和 $R \setminus \overline{G}$ 没有公共点

$$R \setminus \Gamma p(G) = G \cup (R \setminus \overline{G}), G \cap (R \setminus \overline{G}) = 0$$

欲证明圆 K 不是曲线(在乌利松意义下)，只要证明：如果闭集 D 不包含任何由多于一点所构成的连续统，那么它就不能分割圆，因此它就不可能是任何开集 G 的边界. 我们在这里不作这定理的完全的证明，仅仅指定这个证明的需要.

(1) 首先证明：如果列紧集 D 不包含任何连续统，那么对于任意 ε，它都可以表示为有穷个开集的和集的形式，而这些开集的直径小于 ε，并且互无公共点

$$D = D_1 \cup D_2 \cup \cdots \cup D_n$$

$$\delta(D_i) < \varepsilon, i = 1,2,\cdots,n$$

(2) 其次证明：如果 D 是圆的开子集且不包含任何连续统，并且集合 D 分割为闭子集 D_i

$$D = D_1 \cup D_2 \cup \cdots \cup D_n$$

这些闭子集 D_i 的直径小于 ε，而且互无公共点，那么这些子集每个都可以包含





在多角形 P_i 内部,每个 P_i 完全包含在K中,且其直径小于 ε,而且没有两个多角形 P_i 和 P_k 是有公共点的.

(3) 根据(2) 可以找到多角形有限族

$$P_1^1, P_2^1, \cdots, P_{n_1}^1$$

这些多角形互无公共点而其直径小于 1,并且集合 D 包含在这些多角形的和集中,我们管它们叫一级多角形.同样,对于每个一级多角形可以找到直径小于$\frac{1}{2}$的多角形的有限族,它们互无公共点而且完全包含在相当的一级多角形中.这些多角形选取得务使集合 D 包含在所有这些多角形的和集中.这样一来,我们就得到二级多角形的有限族

$$P_1^2, P_2^2, \cdots, P_{n_2}^2$$

这些多角形互无公共点而其直径小于$\frac{1}{2}$,并且集合 D 包含在它们的和集中.

作同样的推论,对于每个自然数 k,我们得到 k 极多角形

$$P_1^k, P_2^k, \cdots, P_{n_k}^k$$

的有限族,这些多角形互无公共点而其直径小于$\frac{1}{k}$,而且它们包含在$(k-1)$级多角形中,并且集合 D 包含在它们的和集里.

(4) 用 F_k 表示所有 k 级多角形的和集,我们便得到闭集的一个降序列

$$F_1 \supset F_2 \supset \cdots \supset F_k \supset \cdots$$

集合 D 便是它们的交集.

(5) 从圆中减去每个集合 F_k,我们就得到连通开集 G_k 的一个升序列

$$G_1 \subset G_2 \subset \cdots \subset G_k \subset \cdots$$

除去集合 D 的点以外,圆的所有点都出现于它们的和集中,作为具有公共点的连通开集 G_k 的和集,集合 S 是连通集.

这样一来就证明了不包含任何连续统的有界闭集 D 不分割圆①.由此推出:圆不是乌利松意义上的曲线,也就推出原先的断言:所有乌利松意义上的平面曲线都是康托意义上的曲线.

4.2 支指数及例

乌利松所给出的曲线定义不仅彻底地阐明了这个概念的本质,而且指出了研究曲线一般性质所必须遵循的道路.连续统的一维性是乌利松定义的要点,也就是连续统 C 的任意点总具有任意小的邻域,其边界不包含任何由多于一

① 同样可以证明,不包含任何连续统的有界闭集 D 不分割平面.这个命题便是所谓菲拉格曼-布劳威尔(Phragman-Browwer)定理的内容.

点所构成的连续统.在这情况,一维的连续统就叫做曲线.在连续统 C 的点的邻域中,只有其边界包含“最少量的点”的邻域是有兴趣的,此处“在指定情况下的点的量”这句话应该了解为边界点集合的势.

设 C 是一曲线,如果对于所有 ε,都可以找到直径小于 ε 的开集(对于 C),它包含 x 点,而其边界包含不多于[①] m 个点,并且对于充分小的 ε,就不能找到直径小于 ε 的开集,它包含 x 点而其边界包含少于 m 个点,那么我们就说,曲线 C 在 x 点具有支指数 m.如果曲线 C 的点 x 具有支指数 m,那么我们就写

$$\mathrm{ind}_x C = m$$

但是开集的边界恒为闭集合,而列紧集中的所有闭集可能是有穷的,或是可列的,也可能具有连续统的势[②].因此,所有开集(对于列紧集)的边界可能是由有穷个点所构成,也可能是可列集,或具有连续统的势.因此,我们按照支指数的观点把曲线的点分类如下:

(1) 设 x 是曲线 C 的一点,n 是一自然数,如果对于一切的 ε,总可以找到(直径小于 ε 的)开集,它包含着 x 点,而其边界包含不多于 n 个点;但对于充分小的 ε,所有包含 x 点而直径小于 ε 的开集的边界都不可能由少于 n 个点所构成,于是我们就说,曲线 C 的 x 点具有有穷的支指数 n.

(2) 设 x 是曲线 C 的一点,如果对于任意 ε,总可以找到直径小于 ε 的开集,它包含 x 点,而其边界是由有穷个点所构成;但当 ε 减小时,包含着 x 点而直径小于 ε 的开集边界点的最小量无限制地增大,于是我们就说,曲线 C 的 x 点具有无界的支指数 ω.

(3) 设 x 是曲线 C 的一点,如果对于任意 ε,总可以找到直径小于 ε 的开集,它包含 x 点,而其边界有由不多于可列个点所构成(这就是说它是有穷集或可列集),但对于充分小的 ε,没有一个包含 x 点而直径小于 ε 的开集,它的边界是由有穷点集所构成的,于是我们就说,曲线 C 在 x 点具有可列的支指数 $\aleph_0$.

(4) 设 x 是曲线 C 的一点,如果对于充分小的 ε,没有一个包含 x 点而直径小于 ε 的开集,它的边界是由有穷或可列的点集合所构成,于是我们就说,曲线的 x 点具有连续统的支指数.

曲线 C 上其支指数大于 2 的点叫做曲线的分支点,而支指数等于 1 的点叫做端点.

在前一节中,我们已经看到,如果连续统 K 是曲线 C 的子集,则其本身也是曲线(定理 1);也看到,与曲线同胚的集合也是曲线(定理 2).下列的两个定理使前面的命题更确切一些,因为这两个定理的证明很简单,所以我们留给读者自己去证.

---

① 在势的意义之下.
② 参看亚历山大洛夫著的《集与函数的泛论初阶》.





定理 1 如果连续统 K 是曲线 C 的子集，x 是连续统 K 的一点，那么 x 点关于 K 的支指数不大于 x 点关于 C 的支指数

$$\operatorname{ind}_x K \leqslant \operatorname{ind}_x C$$

定理 2 如果 C 和 C' 是互相同胚的曲线，x 是曲线 C 的任一点，而 x' 是曲线 C' 在指定的拓扑映象下对应于 x 点的一点，那么点 x 和 x' 的支指数相等

$$\operatorname{ind}_x C \leqslant \operatorname{ind}_{x'} C'$$

欲了解这许多牵涉支指数概念的定义，让我们来考虑一系列阐明这个概念的例子.

例 1 线段在其一切内点皆具有支指数 2，而端点的支指数等于 1.

事实上，一方面，线段的每个内点都可以用长度任意小的区间包围起来，而区间是一个开集(对于线段)，其边界是由两点(区间的端点) 所构成，因此线段的每个内点的支指数小于等于 2.

另一方面，命 G 是任意包含 x 点的开集，如果开集 G 的直径如此之小，以至线段的端点不属于集合 G，那么这对于线段的开集 G 将对于全直线亦为开集.

但是直线上的一切开集都是由区间的有穷或无穷集合所构成，而这些区间互无公共点，并且点 x 属于其中之一. 包含着 x 点的区间的端点是在开集 G 的边界上，因此包含着 x 点的开集的边界是由不少于两点所构成.

这样一来，线段的所有点包含在直径任意小的开集中，其边界是由两点构成，而同时，对于充分小的 ε(小于点 x 到线段端点的距离的一半)，所有包含 x 点而直径小于 ε 的开集的边界至少包含两点. 所以线段的每个内点的支指数等于 2.

线段 ab 的端点 a 可以包含在长度任意小的半区间 $[a, x)$ 之内；这半区间将是线段上的开集，其边界是由一点 x 所构成. 另一方面，每个包含 a 点的开集，只要它的直径小于线段之长，它总包含着半区间 $[a, x)$，因此线段的端点的支指数等于 1.

例 2 圆周的每一点具有支指数 2.

诚然，一方面，圆周的每一点 x 都可以包在一个长度任意小的“开”弧(无端点) 中. 这种弧便是圆周上的开集，这开集的边界是由两点(弧的端点) 所构成. 因此，在圆周每一点的支指数小于等于 2. 另一方面，圆周上的每一开集都是由开弧的有穷或无穷(可列) 集合所构成，这些开弧互无公共点而其端点是在这开集的边界上. 这些开弧中有一个开弧包含着 x 点. 因此，圆周每一点 x 的支指数大于等于 2. 比较二不等式，我们就得到结论：圆周每一点 x 的支指数等于 2.

按语 由于本节的定理 2，在拓扑映象下，点的支指数不变，那么再根据刚才所得到的结果，我们可以断言：

(1) 单弧在它每个内点的支指数等于 2，而单弧端点的支指数等于 1.

(2) 单闭曲线的每个点的支指数等于 2.

例 3 曲线 L 是由从点 o 引出的 n 条直线线段

$$oa_1, oa_2, \cdots, oa_n$$

所构成(图 19.12).

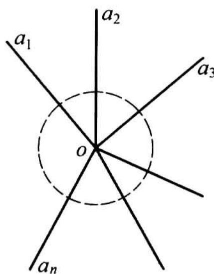


图 19.12

曲线 L 在点 o 的支指数等于 n.

诚然,命 ε 是任一正数.让我们来考虑 n 个半区间

$$oc_1, oc_2, \cdots, oc_n$$

这些半区间是在已给的线段上,而其长度等于$\frac{\varepsilon}{2}$.这些半区间的总体是曲线 L 的开集合,它包含着 o 点而其边界由 n 个点(这些半区间的端点) 所构成.因此曲线 L 在点o 的支指数小于等于n.另一方面,如果数 ε 比所有线段$oa_1, oa_2, \cdots, oa_n$ 的长度都小,那么所有包含点 o 的开集,都包含着 n 个由这一点出发的半区间.这些半区间的端点是在所考虑的开集的边界上.因此曲线 L 在点 o 的支指数大于等于 n.

比较所获得的两个不等式,我们就得到结论:曲线 L 在点 o 的支指数等于 n.在线段 $oa_1, oa_2, \cdots, oa_n$ 的内点,曲线 L 的支指数等于 2,而在线段的端点,L 的支指数等于 1.

例 4 曲线 L 是由引自坐标原点的线段的可列集

$$oa_1, oa_2, \cdots, oa_n, \cdots$$

所构成(图 19.13).这些线段的长度

$$oa_1 = \frac{1}{2}, oa_2 = \frac{1}{4}, \cdots, oa_n = \frac{1}{2^n}, \cdots$$

而这些线段和 ox 轴所构成的角度分别是

$$\frac{\pi}{2}, \frac{\pi}{4}, \cdots, \frac{\pi}{2^n}, \cdots$$

这样一来,线段的长度以及这些线段和 ox 轴正方向所构成的角度将随 n 的增大而趋于零.



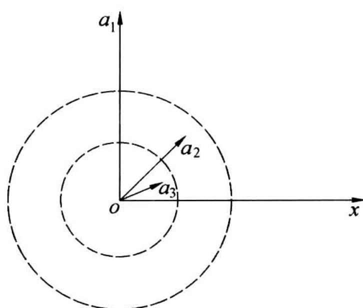


图 19.13

曲线 L 的点 o 具有无限的支指数 ω,也就是说这一点 o 具有以下的性质:尽管对于任意的 ε,都存在着包含点 o 而直径小于 ε 的开集,而这开集的边界是由有穷个点所构成,但是当 ε 趋于零时,这些点的个数将无限增大.

事实上,点 o 的 $\frac{\varepsilon}{2}$ - 邻域是由长度大于 $\frac{\varepsilon}{2}$ 的那些线段上引自点 o 的有穷个半区间以及其长度小于 $\frac{\varepsilon}{2}$ 的无穷个线段所构成. 这个开集的边界是由有穷个点 —— 邻域构造中的半区间的端点所构成,而它的直径小于 ε.

另一方面,无论曲线 L 的开集如何,只要包含点 o 而直径小于 1,那么它包含点 o 的连通区 *компонента*①是由有穷个引自点 o 的半区间以及无穷个引自点 o 的线段所构成的开集. 这连通区的边界是由半区间的端点所构成,并且是在整个开集的边界的成分中. 开集的直径越小,因而它包含点 o 连通区的直径也就越小,那么进入这个连通区中的半区间就越多,因而在这开集的边界上出现的点 —— 半区间的端点就越多. 当开集的直径趋于零时,这些点的数目无限增大.

在构成曲线 L 的诸线段的一切内点,支指数等于2;而这些线段的端点的支指数等于 1.

例 5 曲线 L(图 19.14) 是由长度是 1 的线段 oa_0 以及长度是 1 的无穷个线段

$$oa_1, oa_2, \cdots, oa_n, \cdots$$

所构成,这些线段都是从点 o 引出而且和线段 oa_0 所形成的角度分别等于

$$\frac{\pi}{2}, \frac{\pi}{4}, \cdots, \frac{\pi}{2^n}, \cdots$$

那么曲线在线段 oa_0 的每一点都具有可列的支指数.

---

① 包含着 x 点而包含在集合 R 中的最大的连通集 K 叫做 x 点在集合 R 中的连通区. "最大" 的意思是,任何集合 M,一旦包含 K 作为其子集,那么它就不是连通集.

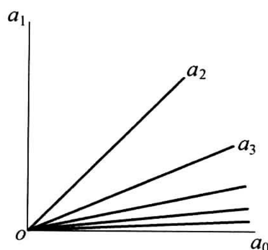


图 19.14

诚然,无论包含着点 o 的开集怎样,它都是由线段 $oa_0, oa_1, oa_2, \cdots$ 上的有穷个或可列个区间(可以根本没有这些区间) 以及引自点 o 的区间的可列集所构成的点 o 的连通区所形成.

如果 m 是线段 oa_0 的一个内点,那么任何包含此点的开集是由所有线段 $oa_0, oa_1, oa_2, \cdots$ 上的区间所构成,当然在这些线段之中也可能除去有穷个线段

$$oa_1, oa_2, \cdots, oa_n$$

最后,所以包含着 a_0 点的充分小的开集都是由其端点是在

$$a_0, a_n, a_{n+1}, a_{n+2}, \cdots$$

(属于这开集) 的半区间的可列集以及线段

$$oa_0, oa_n, oa_{n+1}, \cdots$$

上的有穷个或可列个(可能完全没有) 区间所构成.

因此,我们看到,在所有情况下,包含线段 oa_0 的点的开集的边界恒由点的可列集所构成.所以线段 oa_0 的所有点具有可列的支指数.而在曲线 L 的其他的点,支指数等于 2 或 1.

例 6 曲线 L 是由线段序列(参看图 19.7)

$$A_1, B_1; A_2, B_2; \cdots; A_n, B_n; \cdots$$

以及极限线段 A_0 所构成,其中 A_n 是由诸点

$$x = \frac{1}{2^n}, -1 \leqslant y < 1$$

所构成的线段,B_n 是由诸点

$$\frac{1}{2^{n+1}} \leqslant x \leqslant \frac{1}{2^n}, y = (-1)^n$$

所构成的线段.而极限线段 A_0 是由诸点

$$x = 0, -1 \leqslant y \leqslant 1$$

所构成.

我们在 2.4 末就提到过,这集合同胚于由函数



$$y = \sin\frac{1}{x}, 0 < x \leqslant 1$$

的圆形的点以及极限线段

$$x = 0, -1 \leqslant y \leqslant 1$$

的点所构成的集合.

我们首先证明,集合 L 在我们定义的意义之下确实是曲线.这集合是列紧集.事实上,命 M 是包含在 L 中的无穷子集.如果在集合 M 中具有点的无穷子集 N,N 的点是属于线段

$$A_0, A_1; B_1, A_2; B_2, \cdots$$

中的某一线段,那么由于线段的列紧性,集合 N 将具有属于这同一线段的极限点 p.点 p 既然是集合 N 的极限点,更加是集合 M 的极限点.如果集合 M 包含无穷个不同线段 $A_1, B_1, A_2, B_2, \cdots$ 的各点,那么在线段 A_0 必定可以找到一点,它是集合 M 的极限点.

集合 L 是连通的.事实上,由线段

$$A_1, B_1; A_2, B_2; \cdots; A_n, B_n, \cdots$$

所构成的集合 S 是连通的,因为此序列的相邻二线段具有公共点.把线段 A_0 的点(集合 S 的极限点)添到集合 S 中去,我们就得到连通集 L.

集合 L 既然是列紧的而且是连通的,所以是连续统.我们来证明连续统 L 的曲线.如果点 m 属于线段 $A_n, B_n (n \neq 0)$ 的某一个,那么每个这种点的支指数等于 2,但要除去点$(\frac{1}{2}, 1)$(线段 A_1 的端点),这一点的支指数等于 1.

命 m_0 是极限线段 A_0 的一点,ε 是任一正数,它小于从 m_0 点到极限线段的每个端点的距离.我们用 U 来表示由线段 A_0 及线段

$$A_n, A_{n+1}, A_{n+2}, \cdots$$

$(\frac{1}{n} < \frac{\varepsilon}{2})$ 上的区间所构成的开集,所有这些区间的长度等于 ε,而且其中点是在通过 m_0 点而平行于 ox 轴的直线上.此开集包含 m_0 点,其直径小于 ε,而其边界是由点的可列集所构成,这些点就是这开集的构成区间的端点.因此,开集合 U 的边界不包含任何由多于一点所形成的连续统.

这样一来,我们就证明了集合 L 是曲线.我们证明这曲线在线段 A_0 的每点具有可列的支指数.

事实上,所有包含极限线段 A_0 的 m_0 点的开集,包含着线段 A_0 上及所有垂直线段

$$A_n, A_{n+1}, A_{n+2}, \cdots$$

(从某个号码 n 开始) 上的区间,而且这些区间的端点是在所考虑的开集的边界上.因此,包含着极限线段 A_0 的 m_0 点的开集的边界总是由点的可列集所构

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

成,所以,m_0 点具有可列的支指数 $\aleph_0$.

例 7 曲线 L 是由连接点 o 与康托完全集 C 的点的线段所构成(图 19.15).

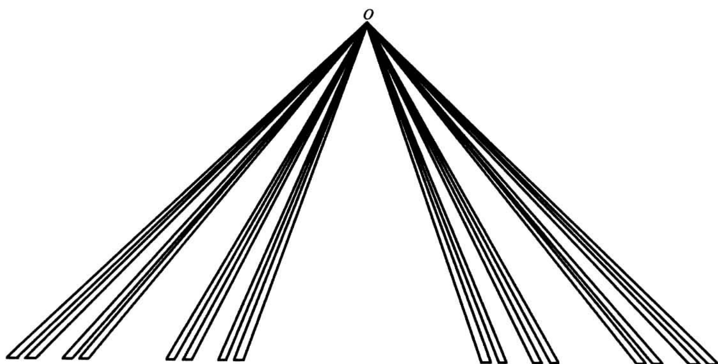


图 19.15

我们首先来证明这个集合是曲线.我们留给读者自己去证明 L 是连续统.

我们来证明,对于任意正数 ε,连续统 L 的每个点都包含在直径小于 ε 的开集中,此开集的边界不包含任何由多于一点所构成的连续统.欲验证此事,只要取一个梯形,此梯形包含已给点在其内部,其底边平行于水平线段,而其侧边是在连接着点 o 和相邻区间的点的线段上.于是,梯形应该取得充分小,以使其边长和对角线的长小于 ε.连续统 L 在这梯形内部的点的总体将是开集,其边界不包含任何由多于一点所构成的连续统,因为这边界是由两个康托完全集所构成,这二安全集就是梯形的底边和连续统 L 相交而得到的.如果点属于集合 C 或者和点 o 重合,开集可以同样地被建立.这样一来,我们就证明了连续统 L 是曲线.

现在我们来证明,曲线 L 的每一点都具有连续统的支指数.我们知道,所有包含在 R 中的开集的边界分割集合 R,这意思是 $R \setminus \Gamma p(G)$ 可以表示为无公共点的二开集 G 和 $R \setminus \overline{G}$ 的和集形式.因此,欲证明曲线 L 的每个点都具有连续统的支指数,应该证明无论对于怎样的开集,其边界不可能由点的有穷集或可列集所构成.为此,也只要证明:如果从曲线 L 中选出它的点的有穷或可列集 A,那么此后所剩下的集合 $L \setminus A$ 将是连通集.

命 A 是不包含点 o 的可列集.不包含集合 A 的点的那些线段的点的集合 S 是连通集.连续统 L 的所有点 x,在其邻域中包含着属于不可列个(构成曲线 L)的线段的点,因此集合

$$B = L \setminus A$$

的所有点 x 将都是连通集 S 的极限点.把其极限点添入连通集 S,我们仍然得到连通集.因此集合 $B = L \setminus A$ 是连通集,所以集合 A 不可能作为曲线 L 的任何开



集的边界.

例 8 希尔宾斯基覆盖 S 在它每一点都具有连续统的支指数.

和上面一样,欲验明此事,只要证明:如果我们从连续统 S 中取出可列闭集 F,则剩下的集合

$$G = S \setminus F$$

是连通集.

假定相反,那就是说开集 $G = S \setminus F$ 不是连通的,于是它就可以表示为两个不相交的非空开集 G_1 与 G_2 的和集的形式

$$G = G_1 \cup G_2$$

我们用 Q 表示单位正方形,希尔宾斯基覆盖就是从这个正方形中得来的,而用 C 表示这个正方形的周界;我们首先来证明,正方形 Q 的周界 C 完全属于 G_1 或 G_2 的一个集合的闭包中,我们用 P 来表示适合下面性质的线段所构成的集合,这些线段都包含在 S 中,并且平行于正方形的边,而不包含集合 F 的点,并且这些线段的长度是 1. 因为每个这样的线段都是连通集,所以如果它包含在两个不相交的开集的和集中,那么它就不可能和这两个开集都有公共点. 我们用 D_1 表示这些平行于正方形之水平边而包含在 G_1 中的线段之一. 于是每个不包含集合 F 的点的垂直线段也将完全属于集合 G_1,这是因为它和这集合 G 有公共点(即它和水平线段 D_1 的交点) 的缘故. 但在这情况下,所有不包含集合 F 的点的水平线段 D 和所有垂直线段(属于 G_1 中) 皆相交,因此将完全包含在 G_1 中. 这样一来,我们就证明了由不含集合 F 的点的水平和垂直线段所构成的集合 P 完全包含在 G_1 中. 周界 C 的所有点都是集合 P 的极限点,因此 $C \subset \overline{G_1}$.

我们用 Q^n 表示一个 n 级正方形,而且 C^n 表示这正方形的周界. 因为曲线 S 在 Q^n 中的部分和整个曲线具有相同的结构,所以同以上一样,可以证明,对于任意 n,每个 n 级正方形 Q^n 的周界 C^n 都属于集合 $\overline{G_1}$ 或 $\overline{G_2}$ 之一.

但如果 $(n+1)$ 级正方形 Q^{n+1} 包含在 n 级正方形 Q^n 中,那么它的周界 C^{n+1} 同正方形 Q^n 的周界具有公共点. 又因为这些公共点的集合是由一个或两个线段 —— 较小的正方形的边所构成,因此是不可列的集合,那么由此推出: C^{n+1} 和 C^n 具有不属于 F 的公共点. 因此,周界 C^{n+1} 和 C^n 包含在同一个集合 $\overline{G_1}$ 或 $\overline{G_2}$.

因为,按定义,正方形 Q 的周界 C 是包含 $\overline{G_1}$ 中,那么由此推出:所有一级正方形的周界都包含在 $\overline{G_1}$ 中. 完全一样可以证明:二级正方形以及所有以后各级正方形的周界也都包含在 $\overline{G_1}$ 中. 最后,因为所有级正方形的一切周界的和集在 S 中稠密,那么由此推出:整个集合 S 是包含在 $\overline{G_1}$ 中,因而集合 G_2 是空集. 这样

一来,就证明了:如果 F 是希尔宾斯基覆盖 S 中的可列集,那么集合 $S \setminus F$ 是连通集.

按语 比较一下后面的四个例子,我们便看到具有无穷支指数的点不是孤立地出现的,这并不是偶然的,正如乌利松所指出:它们恒有如此之多,以至每个这种点的任意 ε - 邻域中都包含完全是由支指数无穷的点所构成的连续统.

例 9 曲线 L 是由点 0 以及可列个两两相切的圆周所构成(图 19.16),这些圆周的直径是

$$\left[1,\frac{1}{2}\right],\left[\frac{1}{2},\frac{1}{3}\right],\cdots,\left[\frac{1}{n},\frac{1}{n+1}\right],\cdots$$

因为,如果我们取 0 点的 $\frac{1}{n}$ - 邻域,那么这邻域的边界是由一个点 $\frac{1}{n}$ 所构成,所以曲线 L 在 0 点具有支指数 1. 曲线 L 在诸圆周的接触点具有支指数 4;在其他各点,曲线的支指数等于 2.

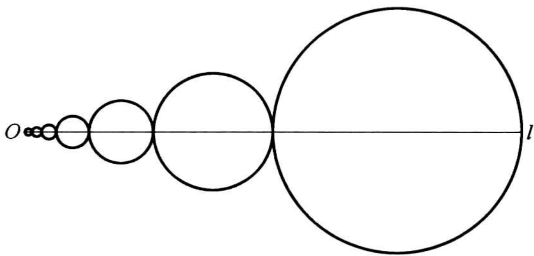


图 19.16

例 10 设既约分数 $\frac{p}{q}$ 是横轴上区间 $0 \leqslant x \leqslant 1$ 的任一有理点,于此点作横轴之垂直线段,线段位于上半平面而其长度等于 $\frac{1}{q}$. 曲线 L 是由横线段

$$0 \leqslant x \leqslant 1, y = 0$$

以及其有理点上的诸垂直线段所构成(图 19.17).

在各垂直线段的端点,曲线 L 的支指数等于 1;而在这些线段的内点,因为对于充分小的 ε,垂直线段的点的 ε - 邻域中,除去此线段的点以外,不包含任何其他点,所以曲线 L 的支指数等于 2. 横线段上的无理点以及其端点都具有支指数 2. 诚然,无论横线段上的无理点 x_0 及正数 ε 如何,恒可以找到这样的有理正数 ε',使所有界于直线

$$x = x_0 - \varepsilon' \text{ 和 } x = x_0 + \varepsilon'$$

之间的带域内的垂直线段的长度小于 ε. 我们用 ε'' 表示所有界于直线

$$x = x_0 - \varepsilon' \text{ 和 } x = x_0 + \varepsilon'$$



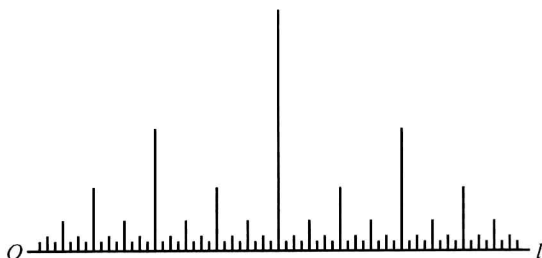


图 19.17

之间的带域内的垂直线段的最大长度，$\varepsilon'' < \varepsilon$. 曲线 L 上适合条件

$$x_0 - \varepsilon' < x < x_0 + \varepsilon'$$

$$0 \leqslant y < \varepsilon''$$

的点所构成的开集包含 x_0 点，而其本身包含在 x_0 点的 ε – 邻域中；其边界是由横线段上具有横坐标 $x_0 - \varepsilon'$ 和 $x_0 + \varepsilon'$ 的两点所构成.

如果 $x = \dfrac{p}{q}$ 是横线段上的一个有理点，那么曲线 L 在 x_0 点的支指数等于 3. 事实上，如果 ε 小于读数 $\dfrac{p}{q}$，$1 - \dfrac{p}{q}$ 和 $\dfrac{1}{q}$ 中的每一个数，那么所有直径小于 ε 而包含 x_0 点的开集之边界皆与引自 x_0 点的三线段相交，所以曲线 L 在 x_0 点的支指数大于等于 3. 另一方面，作同上之推论，总可以找到这样的开集，它包含 x_0 点而包含在此点的 ε – 邻域之中，并且其边界是由引自 x_0 点，的线段上的三点所构成. 由此推出：曲线 L 在 x_0 点的支指数小于等于 3. 比较所获得的二不等式，我们求得，曲线 L 在 x_0 点的支指数等于 3.

例 11 曲线 C 建立如下：在边长等于 1 的等边三角形内，引三条中线（这里的中线是指三角形三边中点之连线而言，与通常的所谓中线不同，切勿混淆 —— 原译者），挖掉这三中线所围成的三角形的内点. 剩下的集合是由三个三角形所构成，我们称这些三角形为一级三角形. 对于每个一级三角形，施行同样的手续：即引它们的中线，且挖掉这些中线所围成的三角形之内点. 以同样的方式来处理这所得的 9 个二级三角形，便得到 27 个三级三角形. 继续这样处理下去，于是对于每个自然数 n，我们就得到由 3^n 个 n 级三角形所构成的集合.（图 19.18）

完成了所有这些手续以后，所剩下的集合 C 是一个连续统. 事实上，这集合 C 可以表示为许多连续统的公共部分，这些连续统中的第一个是由三个一级三角形所构成，第二个是由 9 个二级三角形所构成，第三个是由 27 个三级三角形所构成，如此等等；这些连续统中的每一个都包含在前一个之中. 所以（按 2.6 定理 1），这些连续统的公共部分仍为连续统.

这连续统 C 是曲线. 欲证明这一点，我们验明连续统 C 的点具有支指数 2，

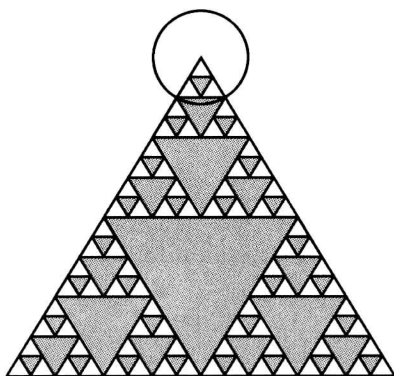


图 19.18

3,或 4.

在基本三角形的顶点,连续统 C 具有支指数2.诚然,一方面对于所有 n,顶点的$\frac{1}{2^n}$ - 邻域的边界是由两点所构成,因此,在基本三角形的顶点,连续统 C 的支指数小于等于 2.另一方面,所有包含基本三角形顶点的充分小的开集之边界,必定和引自此顶点的两边相交,由此推出连续统 C 的支指数大于等于 2.比较这两不等式,我们发觉,连续统 C 在基本三角形的顶点的支指数等于 2.

在分割基本三角形时所得到的各级三角形的顶点,连续统 C 的支指数等于4.命 x 是这样的一个顶点.于是,一方面对于所有 n,点 x 的$\frac{1}{2^n}$ - 邻域之边界是由四点所构成,这四点就是以 x 点为其公共顶点的那两个 n 级三角形的顶点.因此,连续统 C 在 x 点的支指数小于等于 4.

另一方面,设 k 是能使 x 为 k 级三角形顶点的最小的一个数,那么仍有一 k 级三角形存在,x 点为其顶点.因此,有四个线段由 x 点发出,这四个线段就是以 x 点为其公共点的 k 级三角形的边.所有包含 x 点的充分小的开集的边界应该和这些边相交,因此连续统 C 的支指数大于等于 4.比较二不等式,我们便得结论:在所有级三角形的顶点,连续统 C 的支指数等于 4.

最后,如果 x 是连续统 C 的任一点,它既不是基本三角形的顶点,也不是部分三角形的顶点,那么在所有这样的点,连续统 C 的支指数等于 3.当然,对于所有 n,可以找到一个 n 级的三角形,点 x 在此三角形的内部或在它的边上.在所说的包含着 x 点的 n 级三角形的外接圆之内部,连续统 C 的点所构成的开集之边界是由三点所构成,这三点就是所说的三角形的顶点.因为借适当地选取 n,而可以使得这开集的直径任意小,所以由此推出,连续统 C 在 x 点的支指数不大于 3.

另一方面,对于每一点 x,总可以指出一个包含着 x 点的三角形的序列,在



这序列中,每后一三角形都包含在前一三角形中.点 x 是这序列三角形顶点的极限点.序列中每两个相邻的三角形仅有一个公共顶点,而后一三角形的其他二顶点是在前一三角形的边上.这些三角形的交集是 x 点.用这些三角形的边的线段.我们可以作成三个无穷节的折线,点 x 是这三折线的极限点.这些折线是由前一三角形的顶点到后一三角形顶点的边的线段所构成.如此所得到的这些折线没有任何公共点,更因为点 x 是其每个的极限点,所以任意包含着 x 点的开集的边界应该和这些折线每一个相交.因此,连续统 C 在 x 点具有支指数大于等于3.比较二不等式,我们得到结论:连续统 C 在 x 点具有支指数3.

按语1 如果我们考虑由二连续统 C_1 和 C_2 所构成的集合,这二连续统每一个皆和例11中的连续统同胚,而且除了对应于连续统 C 的基本三角形顶点的那些点以外,这二连续统 C_1 和 C_2 没有其他的公共点,那么我们便得到具有支指数3和4的点所构成的曲线的实例.

按语2 只要把上例稍微变化一下,就可以使得曲线 C 只有支指数等于2或3的点.

要达到这个目的,不要取三角形的中线,而连接三角形各边上离顶点距离为边长三分之一的各点,然后挖掉这些连线所围成的正六边形的内点(图19.19).

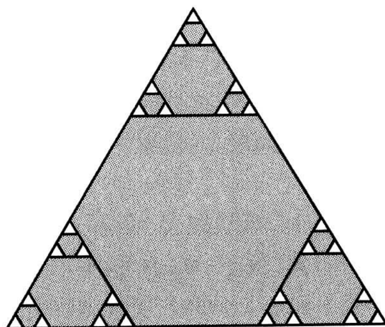


图19.19

4.3 有穷支的曲线

曲线上,支指数大于2的点(具有无界的和无穷的支指数的点也算在内),称为这曲线的分歧点(Точка ветвления),而支指数等于1的点叫做这曲线的端点(Концевая точка).

上面的例子说明:曲线可以随其各式各样的分歧点的存在而呈各种不同的形式.在这里不可能停留在曲线牵涉它的分歧点的所有各种各样的性质,我们只限于考虑仅有有穷个分歧点的那种曲线,而且在它每个分歧点的支指数为有穷.这样的曲线,我们将称之为有穷支的曲线.

在这里,我们要建立许多事实,这些事实将指出:支指数的概念何等深刻地描述着曲线的特征.

如果曲线完全没有分歧点,即曲线在每点的支指数等于1或2,那么这曲线或者是单弧——线段的拓扑影象,或者是单闭曲线——圆周的拓扑影象.于是,如果曲线在所有点的支指数等于2,那么曲线是圆周的拓扑影象(单闭曲

线);如果曲线的端点(可以证明,端点必定是两个),那么这曲线便是单弧.

如果曲线只有有穷个分歧点,而且每个分歧点的支指数也都是有穷,那么这样的曲线可以分割为有穷个单弧,除其端点外,这些单曲线互相没有其他的公共点.

所有这些论断的证明都基于以下的事实:有穷支的曲线的任意两点,皆可用单弧连接之.而最后的这个论断又可以(由于2.6定理3)从下面的事实推出来:有穷支的曲线是局部连通的连续统.从这个命题的说明开始,我们阐明所有上述断言的正确性.

定理 1　如果连通集在其每点都具有有穷的支指数,那么此集合便是局部连通集.

为了证明这个定理,我们需要下面一个简单的引理:

引理 1　如果集合 M 不可能表示为 n 个连续集的和集的形式(也就是说,无论集合 M 怎样表示成 n 个集合的和集的形式,这 n 个集合中至少有一个不是连通集),那么集合 M 就可以表示为 $n+1$ 个互无公共点的闭集(关于 M)的和集的形式.

事实上,如果集合 M 不是连通集,那么按定义,它就可以表示为两个无公共点的闭集(关于 M)的和集形式.因此我们的引理对于 $n=1$ 是正确的.假设对于从1到 $n-1$ 的一切数,我们的引理都已经证明了,我们来证明对于数 n 的情形.

如此,我们假设集合 M 不可能表示成 n 个连通集的和集形式.在这种情形,它就更不能表示为 $n-1$ 个连通集的和集形式.但此时由于归纳法的假定,集合 M 是

$$(n-1)+1=n$$

个互不相交的闭集

$$A_1, A_2, \cdots, A_n$$

的和集.因为集合 M 不能表示为 n 个连通集的和集形式,所以在这些集合 A_1, $A_2, \cdots, A_n$ 中,至少有一个不是连通集.假如命这个是集合 A_n,于是它就可以表示为两个无公共点的非空间集 A_n^* 和 A_{n+1} 的和集形式,这两个集合 A_n^* 和 A_{n+1} 是关于 A_n 的闭集,因而关于 M 亦为闭集.这样一来,整个集合 M 便可以表为 $n+1$ 个互无公共点的闭集

$$A_1, A_2, \cdots, A_n^*, A_{n+1}$$

的和集形式.

我们的引理就被证明了.

现在我们来证明定理1.我们需要证明:无论集合 R 的点 x 以及包含此点 x 的开集 U 如何,总存在一个连通开集 V,这开集 V 包含着 x 点而其本身包含在 U





之内.

因为集合 R 在它每点都具有有穷支指数,所以存在一个包含 x 点而其本身包含在集合 U 中的开集 G,此开集的边界是由有穷个点

$$a_1, a_2, \cdots, a_n$$

所构成.

首先,我们证明集合

$$\overline{G} = G \cup a_1 \cup a_2 \cup \cdots \cup a_n$$

可以表示为 n 个连通集的和集形式.我们假设不是如此,于是根据刚才证明了的引理,集合 $\overline{G}$ 可以表为 $n+1$ 个互不相交的非空闭集(关于 $\overline{G}$ 因而亦关于 R)

$$B_1, B_2, \cdots, B_{n+1}$$

的和集形式.

在闭集

$$B_1, B_2, \cdots, B_{n+1}$$

中没有任何一个可以完全包含在开集 G 中,因为,譬如倘若集合 B_i 包含在 G 中,那么整个集合 R 就可以表示为两个不相交的非空闭集 B_i 和

$$B_1 \cup B_2 \cup \cdots \cup B_{i-1} \cup B_{i+1} \cup \cdots \cup B_{n+1} \cup (R \setminus G)$$

的和集形式.这与集合 R 的连通性相违背.但是既然构成和集 $\overline{G}$ 的诸集合

$$B_1, B_2, \cdots, B_{n+1}$$

中没有一个是完全包含在 G 中的,那么它们每个至少包含集合 G 的边界的一个点,即诸点

$$a_1, a_2, \cdots, a_n$$

之一.但是因为我们一共有 $n+1$ 个集合 B,而只有 n 个点 a,并且没有两个集合 B 是有公共点的,所以那是不可能的.

这个矛盾使我们不得不作出结论说:集合 $\overline{G}$ 不可能表示为 $n+1$ 个互无公共点的非空闭集的和集的形式.由于前面的引理,由此也就推出:集合 $\overline{G}$ 可以表示为不多于 n 个连通集的和集的形式.假定这些加项的数目等于 k.取这些加项的每个集合的闭包,我们便把闭集 $\overline{G}$ 表示为 k 个连通闭集

$$F_1, F_2, \cdots, F_k \qquad ①$$

的和集形式.我们用 C_x 来表示组①中包含着 x 点的那些闭集的和集.点 x 属于集合 G.因此,如果我们从 G 中减去组①中不包含 x 点的那些闭集,那么我们就得到包含着 x 点而其本身包含在 C_x 中的开集 $H(x)$.

这样,我们就证明了:无论集合 R 的点 x 以及包含着它的开集 U 如何,总可以找到这样的连通集 C_x 以及这样的开集 $H(x)$,使得

$$x \in H(x) \subset C_x \subset U$$

我们用 V 表示包含着 x 点而其本身包含在 U 内的最大连通集①. 首先,我们有

$$C_x \subset V \subset U$$

我们证明集合 V 是开集.

命 y 是集合 V 的任意点. 于是按照以前的道理,对于此点 y 可以找到这样的连通集 C_y 以及这样的开集 $H(y)$,使得

$$y \in H(y) \subset C_y \subset U$$

因为 V 和 C_y 具有公共点 y 并且是包含在 U 中的连通集,所以它们的和集

$$V \cup C_y$$

也是包含在 U 中的连通集. 但是,V 是包含着 x 点而其本身包含在 U 中的最大的连通集. 因此

$$C_y \subset V$$

因此,对于连通集合 V 的所有点 y,总可以找到这样的开集 $H(y)$,使得

$$y \in H(y) \subset V$$

因此,集合 V 可以表示为安全包含在 V 中的开集 $H(y)$ 的和集,所以集合 V 本身也是开集.

这样一来,对于一切点 x 以及包含着它的开集 U,我们找到了包含着 x 点而其本身包含在 U 中的连通开集 V,这就证明了集合 R 的局部连通性.

定理2 如果连续统 C 在它的两点 a 和 b 具有支指数 1,而在这连续统其余的所有点,支指数等于 2,那么连续统 C 是单弧 ab.

证明 因为连续统 C 在一切点具有有穷的支指数,所以它是局部连通集. 因此,它的每两点可以用单弧相连接. 命 ab 是连接 a 点和 b 点的单弧. 和定理的命题相反,我们假定连续统 C 包含着一个不属于弧 ab 的点 c. 由于连续统 C 的局部连通性,故可以用单弧 ac 将 c 点与 a 点相连接.

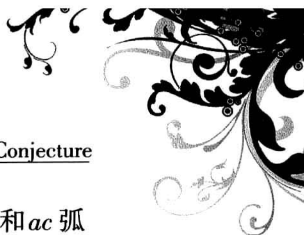
我们用 d 表示 ac 弧上二弧 ac 和 ab 的最后的(从 a 到 c 的方向计算)交点.

如果点 d 不知点 a 或点 b 任一点相重叠,那么在点 d,连续统具有支指数大于等于 3.

事实上,我们取一个包含 d 点的充分小的开集 G,使得点 a,b 和 c 都不属于它. 有三个互无公共点的弧由 d 点出发:弧 da 和 db(合起来构成弧 ab)以及弧 dc. 因为弧 da,db 及 dc 是连接着开集 G 的点 d 和不属于这开集的点 a,b 和 c 的连通集. 那么这些弧应该和开集 G 的边界相交. 因此,所以包含着 d 点的充分小的开集的边界,至少包含三个点,这与连续统 C 的点 d 具有支指数 2 的假设相违背. 如果我们假设点 d 和点 a 或 b 之一相重叠,那么我们看到:这点将具有支指数 2. 譬如,我们假定点 d 和点 a 一致. 于是命 G 是包含着点 d 的充分小的开集,

① V 是所有包含 x 而其本身包含在集合 U 中的连通集的和集.





以至点 b 和点 c 皆不包含在其中,那么开集 G 的边界将既和 ab 弧相交也和 ac 弧相交,因此,按假定和点 a 相一致的点 d 将具有支指数 2,但按照定理的条件连续 C 在点 a 的支指数等于 1,因此和定理的条件相违背.

因此,假定连续统 C 包含着一点 c 不属于弧 ab,在所有情况终归与定理的条件相矛盾.因此连续统 C 与弧 ab 相一致.

定理 3 如果连续统 C 的所有点都具有支指数 2,那么这连续统 C 是单闭曲线.

证明 所有连续统 C 至少包含这样的两个点 a 和 b,这两点不分割这个连续统.这就是说,集合 $C \setminus a$ 和 $C \setminus b$ 每个都是连通集.由于定理所给的条件,连续统 C 的所有点都具有有穷的支指数,所以它是局部连通集.因此,这连续统的任意两个点都可以用单弧来连接.我们用单弧来连接不分割连接统 C 的两点 a 和 b.

因为倘若连续统 C 和单弧 ab 相一致,那么这连续统的点 a 和点 b 便具有支指数 1,这与定理的条件相违背,因此在连续统 C 内总可以找到一点 c 不属于 ab 弧.

因为点 a 不分割连续统 C,也就是说,开集 $C \setminus a$ 是连通集,更由于它的所有点皆具有有穷的支指数,因而此集合为局部连通集,所以这集合的点 b 和 c 可以用一个不包含点 a 的单弧 bc 来连接.完全一样,连通且局部连通的开集 $C \setminus b$ 的两点 a 和 c 也可以用一个不包含点 b 的单弧 ac 来连接.

首先我们证明:弧 ab,bc 和 ca 除其端点以外,不可能有其他的公共点.譬如,我们来证明弧 ca 和 cb 除其端点 c 以外,没有其他任何公共点.假定不是这样,命 d 是弧 ca 上最后的(从 c 到 a 计算)一个属于弧 cb 的点.命 G 是包含着点 d 的任一开集,此集是如此之小,以至诸点 a,b 和 c 都在这开集之外.有三个弧 db,dc 和 da 是从点 d 出发,其中 db 和 dc 是以点 d 分割 bc 弧而得来.这三个弧除去点 d 以外,互无任何其他的公共点.开集 G 的边界与这些弧的每一个皆相交.因此连续统 C 的点 d 具有支指数大于等于 3,这与定理的条件相违背.这个矛盾证明了:除点 c 以外,弧 ca 和 cb 没有任何其他的公共点.

我们证明了除去它们的端点以外,弧 ab,bc 和 ca 没有任何其他的公共点,因此,这三个弧的和集是一单闭曲线 L.我们现在应该证明:除去曲线 L 的点以外,连续统 C 不包含任何其他的点.假定相反,命 d 是连续统 C 不属于曲线 L 的一点.我们把点 d 和曲线 L 的某一点相连接,譬如,用单弧 ad 将点 d 和点 a 相连接,我们用 e 表示 ad 弧上它与曲线 L 最后的(从 a 和 d 计算)交点.运用上面累次采用的方法,我们证明点 e 将具有支指数大于等于 3,这与定理的条件相矛盾.因此,连续统 C 没有任何一点 d 不属于曲线 L.因此,连续统 C 和单闭曲线 L 相一致.

定理 4　如果曲线不包含分歧点,这就是说,此曲线的所有点的支指数小于等于 2,那么这曲线或者是单弧,或者是单闭曲线.

证明　如果不含分歧点的曲线 L 没有端点,即在曲线的所有点,支指数等于 2,那么根据定理 3,曲线 L 是单闭曲线.

如果不含分歧点的曲线 L 具有两个端点,则根据定理 2,这曲线是单弧.因此欲完成这个定理的证明,我们必须证明:如果曲线 L 不含分歧点,那么它或者安全没有端点,或者仅有两个端点.

欲验明这一点,我们应该证明下列二命题:

(1) 如果曲线 L 只有一个端点,那么它至少具有一个分歧点.

(2) 如果曲线 L 的端点的个数大于 2,那么在曲线上必至少有一个分歧点.

我们来证明命题(1).假定曲线 L 只有一个端点 a.因为 L 是连续统,那么按照已经证明了的事实①,在这连续统 L 上至少有两点不分割此连续统 L.

我们用 b 来表示曲线 L 上异于点 a 而不分割曲线 L 的一个点.与命题(1)相反,我们假设曲线 L 没有分歧点,即曲线 L 在它所有点的支指数小于等于 2.于是曲线 L 将是局部连通集,因此它的所有两点皆可以用单弧来连接.

我们用单弧 ab 来连接点 a 和点 b.因为,如果曲线 L 和弧 ab 相一致,那么曲线 L 将有两个端点 a 和 b,这与假设相矛盾,因此在曲线 L 上有一点不属于弧 ab.

命 c 是曲线 L 上不属于 ab 弧的一点.因为点 b 不分割曲线 L,即集合 $L \setminus b$ 是连通(且局部连通)集,故可以用单弧将点 a 与点 c 相连接,而此单弧安全是在开集 $L \setminus b$ 内,因而不包含点 b.我们用 d 来表示 ac 弧上最后的(从 a 到 c 计算)一个它与弧 ab 相交之点.

点 d 不可能与点 a 相一致,因为在此情形下点 a 的支指数将等于 2,这与点 a 是端点的假定相抵触,因此,点 d 异于点 a.在这个情况下,曲线 L 在点 d 将具有支指数大于等于 3,因为,任何包含点 d 的充分小的开集的边界,至少与曲线 L 有三个公共点:其中两点是在 da 弧和 db 弧上(此二弧合起来构成 ab 弧),第三点是在 dc 弧上.所以 d 是曲线 L 的分歧点.这个矛盾就证明了我们的命题(1).

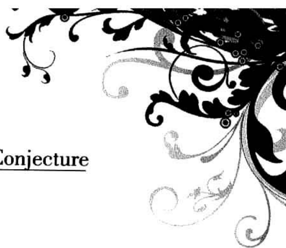
为了证明命题(2),我们需要下面的引理:

引理 2　端点不分割曲线,这就是说,由曲线去掉它的端点而得到的集合是连通集.

命 a 是曲线 L 的一个端点.假设开集(关于 L)$L \setminus a$ 不是连通集,于是它就可以表示为两个无公共点的非空间集 G 与 H 的和集形式.

① 2.6,定理 3.





闭集

$$\overline{G} = G \cup a$$

和

$$\overline{H} = H \cup a$$

每个都是连通集.诚然,譬如倘若我们假设集合 $\overline{G} = G \cup a$ 不是连通集,那么它可以表示为两个无公共点的非空闭集 P 和 Q 的和集形式.我们假定点 a 属于集合 Q.闭集 P 和 $Q \cup \overline{H}$ 没有公共点,且其和集构成 L.而这与 L 的连通性相矛盾,因此集合 $\overline{G}$ 是连通集.完全一样,可以证明集合 $\overline{H}$ 的连通性.

命 U 是任何包含着点 a 的一个开集,此开集如此之小,以至在它之外既有集合 $\overline{G}$ 的点,也有集合 $\overline{H}$ 的点.因为集合 $\overline{G}$ 和 $\overline{H}$ 都是连通集,而且除点 a 以外没有其他的公共点,所以开集 U 的边界应该既与 $\overline{G}$ 有公共点,也与 $\overline{H}$ 有公共点.但是这二集合只有一个公共点 a,因此集合 U 的边界与集合 $\overline{G}$ 和 $\overline{H}$ 相交于两个不同的点.因为按照假定曲线 L 在点 a 的支指数是等于1,因此我们就得到了矛盾.由此推出:集合 $L \setminus a$ 是连通集.我们的引理就证明了.

现在我们来证明命题(2).

设 a,b,c 是曲线 L 的三个端点.其中没有一个分割曲线 L,这就是说,集合

$$L \setminus a, L \setminus b, L \setminus c$$

的每一个都是连通集.

如果我们假定曲线 L 没有分歧点,那么 L 便是局部连通集,因此 a,b 和 c 中的每两点都可以用一个不含第三点的单弧来连接.在三个弧 ab,ac 和 bc 中,每两个皆有异于其端点的公共点,因为譬如倘若弧 ab 和 bc 只有一个公共点 b,那么曲线 L 在此点将有支指数2,这与点 b 是端点的假定相矛盾.因此弧 ab 和 bc 具有异于点 b 的公共点.设 d 是 bc 弧上最后一个属于弧 ab 的点.于是,和以前一样,我们证明曲线 L 在点 d 将有支指数大于等于3,这就是说,d 是分歧点.这个矛盾就证明了命题(2).

在支指数是有穷的情形,下面的定理说明了支指数的几何意义.

定理 5 如果曲线 L 具有有穷个分歧点,而且每个这种点的支指数为有穷,那么从支指数为 n 的分歧点 o 有 n 个单弧引出,除点 o 以外,这 n 个单弧互无其他公共点,而且点 o 的所有邻域,只要它除点 o 以外不包含任何其他分歧点,那么此邻域是由这 n 个弧的点所构成.

证明 由于定理1,曲线 L 是局部连通集,而且因为在点 o 具有支指数 n,所以有这样的连通开集 G 存在,此开集 G 的闭包除点 o 以外不含任何其他的分歧点,而其边界包含着不少于 n 个点.

命

$$a_1, a_2, \cdots, a_n$$

是集合 G 的边界点.

把它的边界添入集合 G,我们就得到局部连通的连续统 $\overline{G}$(参看定理1).此连续统 $\overline{G}$ 包含点 o,对于这连续统 $\overline{G}$ 而言,所有形状如

$$U_i = \overline{G} \setminus (a_1 \cup a_2 \cup \cdots \cup a_{i-1} \cup a_{i+1} \cup \cdots \cup a_n), i = 1,2,\cdots,n$$

的集合皆为连通开集.由于连续统 $\overline{G}$ 的局部连通性,总可以用单弧 oa_i 将点 o 和每个点 $a_i(i = 1,2,\cdots,n)$ 相连接,而这些单弧 oa_i 完全包含在开集(对于 $\overline{G}$)U_i 之内,因此不含

$$a_1, a_2, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n, i = 1,2,\cdots,n$$

中的任何一点.

除去点 o 以外,设弧 $oa_i(i = 1,2,\cdots,n)$ 互无其他的公共点.事实上,如设弧 oa_1 和 oa_2 具有异于点 o 的公共点,那么这不可能是点 a_1,也不可能是点 a_2.如果用 b 来表示 oa_1 弧上同时属于 oa_2 弧的最后的(从 o 到 a_1 算起)一个点,那么此点将具有支指数3,因为从这一点有三个互无公共点的弧引出:弧 bo, ba_2 及弧 ba_1,其中弧 bo 和 ba_2 合起来构成弧 oa_2.但这是不可能的,因为点 b 属于集合 $\overline{G}$,而按照假定,集合 $\overline{G}$ 没有任何异于点 o 的分歧点.

这样一来,我们就证明了至少有 n 个单弧从点 o 出发,这些单弧除了点 o 以外,互无任何其他的公共点.现在我们来证明:开集 G 的每一点属于这些单弧之一.假定相反,即有一点 c 属于集合 G 而不在

$$oa_1, oa_2, \cdots, oa_n$$

中的任一弧上.于是可以用单弧 oc 将点 c 与点 o 相连接.和上面一样,可以证明除去点 o 以外,oc 弧与 $oa_1, oa_2, \cdots, oa_n$ 中的任一弧没有其他的公共点.

我们取点 o 的 ε - 邻域,这样地来选取 ε,使得此数小于从点 o 到各点 a_1, $a_2, \cdots, a_n, c$ 的距离.于是每个包含点 o 而其本身包含在这点的上述 ε - 邻域内的开集的边界,至少包含 $n + 1$ 个点,这是因为每个弧

$$oa_1, oa_2, \cdots, oa_n, oc$$

有一点在开集 G 之内,而有一点在 G 之外,因此皆与此集合的边界相交.但这是不可能的,因为曲线 L 在点 o 具有支指数 n.所得到的这个矛盾证明开集 G 的一切点 c 属于各弧 $oa_1, oa_2, \cdots, oa_n$ 之一.我们的定理就完全证明了.

按语 特别,从这被证明了的定理推出:如果点 o 具有支指数2,那么就有两个除点 o 以外无任何其他公共点的单弧由此点出发,并且点 o 的充分小的邻域是一开弧(无端点),而点 o 为其内点.

如果 o 是端点,那么有唯一的单弧从 o 出发,此单弧以点 o 为其端点,并且点 o 的充分小的邻域是一"半开"弧,点 o 是这弧的端点之一.

一切有穷支的曲线的结构被以下的定理所阐明:

定理6 欲使曲线 L 可以表示为有穷个单弧的和集形式,而这些单弧除其端点以外,互无其他任何公共点,其充分且必要的条件是:这曲线仅有有穷个分





歧点,而且每个这种分歧点的支指数应该是有穷的.

或者自己去证明条件的必要性,我们只来证明它的充分性,而这个证明是以下面的引理为依据的.

引理 3　如果曲线 L 只有有穷个分歧点.而每个分歧点的支指数又是有穷的,那么曲线端点的集合是有穷的.

假定相反,即端点的集合是无穷集.因为曲线 L 是局部连通集,所以它的所有点 o 皆可以用单弧与诸端点相连接.如果在这些弧中,有无穷个除点 o 无其他任何公共点的单弧,那么点 o 的支指数将为无穷或无限.倘若我们假设,有无穷个除点 o 还有另外公共点的弧偶存在,由此将推出:在这些弧的每个上至少有一个分歧点,于是在曲线 L 上将有无穷个分歧点,这与引理的条件相违背.

我们现在来证明定理 6.首先我们证明:从每个支指数等于 2 的点 o 有两个弧引出,这二弧除点 o 以外无其他任何公共点,而其端点或者是曲线 L 的端点,或者是它的分歧点.在后一情况,从点 o 引出的二弧的端点可以是同一个分歧点.在这种情况下,这两个弧合起来便构成一个单闭曲线(图 19.20).

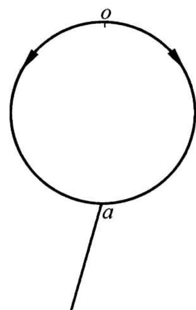


图 19.20

根据前一定理末尾所作的按语,从支指数等于 2 的点 o 引出二个弧,此二弧除点 o 以外无其他公共点,并且不含曲线 L 的分歧点在其内部.命 oa 和 ob 是具有这种性质的最大弧.于是这些弧的端点不可能是支指数 2 的点.事实上,譬如设点 a 具有支指数 2,那么它便可以界于开弧 cd 之内,此开弧便是点 a 的不含分歧点的邻域.此弧端点之一,如点 c,将属于 oa 弧,而另一端点 d 便不属于 oa.除去点 a 以外,弧 oa 和 ad 没有其他任何公共点.因此 od 也是一个单弧,它不含曲线 L 的分歧点在其内部.这样一来,弧 oa 便构成弧 od 的一部分,因此 oa 弧不是不含分歧点在其内部的最大弧.由这矛盾,我们应该断定:曲线 L 在点 a 的支指数不可能等于 2,这就是说,a 或者是端点,或者是分歧点.

如果我们假设,从点 o 出发而不含分歧点在其内部的最大弧的二端点 a 和 b 是曲线 L 的端点,那么在这情形,曲线 L 便是单弧 ab.

事实上,如果集合 $L \setminus ab$ 不是空集,那么和弧 ab 一样,它也是闭集,因为 ab 弧中任意点的充分小的邻域仅包含这个弧的点.所以在这种情况下,曲线 L 可以表示为两个无公共点的非空闭集 ab 和 $L \setminus ab$ 的和集形式,所以不是连通集.

这样一来我们就证明了曲线 L 的所有点 o 皆属于一个单弧,这单弧的两端或者是分歧点(这两分歧点可能重合),或者是一个分歧点和一个端点.但是这

样的弧只可能有有穷个,因为按照我们的假定,曲线 L 只可能有有穷个分歧点,而按照刚才证明的引理,曲线 L 只可能有有穷个端点.但是根据定理 5,只有有穷个弧从每个分歧点引出,而从一个端点只有一个弧引出,因此所有弧的个数是有穷的.因为每个这种弧的内点的支指数等于 2,所以除其端点以外,这些弧中没有两个弧能有其他的公共点.

我们已经看到,圆周(一般所有单闭曲线也是一样)具有这样的性质,即在它的一切点,支指数取同一数值(等于 2).问题就发生了,是否有另一种曲线存在,其所有点的支指数都取同一数值(不等于 2)呢?哪知道无论对于任何有穷数 n,这都不行,因为下列的定理成立.

定理 7 如果曲线 L 的一切点具有支指数大于等于 n,那么在曲线上可以找到一点,其支指数大于等于 $2n-2$.

我们将证明下列形式的定理:如果曲线 L 的一切点具有支指数小于 $2n-2$,那么在曲线上就有一点,其支指数小于 n.

证明 设 A_0 是曲线 L 的任一开集,其直径小于 δ_0.因为曲线在每点的支指数小于 $2n-2$,那么有一开集 G 存在,此开集及其闭包皆包含于 A_0 内,而其边界是由少于 $2n-2$ 个点所构成.命 a_0 是集合 G 的边界点之一.

集合 G 及其闭包都在 A_0 内.因此集合 G 的边界,特别是点 a_0,便属于集合 A_0;因为 A_0 是开集合,所以有这样的正数 ε 存在,使点 a_0 的 ε – 邻域安全包含在集合 A_0 中.此外,我们假设数 ε 既小于 $\frac{1}{2}\delta_0$(这里 δ_0 是集合 A_0 的直径),也小于点 a_0 到集合 G 的其余边界距离之半.

因为曲线 L 在它的所有点具有支指数小于 $2n-2$,特别它在点 a_0 的支指数也小于 $2n-2$,那么有开集 U 存在,此开集 U 包含着点 a 而其本身包含于此点的 $\frac{\varepsilon}{2}$ – 邻域内(此处的 ε 是按照上述的条件所选择的),而它的边界包含少于 $2n-2$ 个点.命 k 是集 U 的边界点的个数,于是

$$k < 2n-2$$

集合 U 及其闭包皆包含在 A_0 中.诚然,无论集合 U 的 x 点如何

$$\rho(a_0, x) \leqslant \frac{\varepsilon}{2} < \varepsilon$$

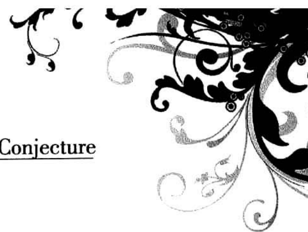
因此点 x 属于点 a_0 的 ε – 邻域,这就是说,它属于集合 A_0.

集合 U 的直径小于 $\frac{\delta_0}{2}$.事实上,如果 x 和 y 是集合 U 的两点,那么

$$\rho(x, y) \leqslant \rho(a_0, x) + \rho(a_0, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon < \frac{1}{2}\delta_0$$

由此推出集合 U 的直径





$$\delta(U) \leqslant \varepsilon < \frac{1}{2}\delta_0$$

除去点 a_0 以外,集合 U 的闭包 $\overline{U}$ 不含集合 G 任一边界点,这是因为,如果 x 是集合 $\overline{U}$ 的一点,那么

$$\rho(a_0, x) \leqslant \varepsilon$$

而 ε 选取得使它小于点 a_0 到集合 G 的其余边界点距离之半.

我们用 H 表示集合 $\overline{G}$ 的补集(它是一开集)

$$H = L \setminus \overline{G}$$

我们考虑开集

$$U \cap G \text{ 和 } U \cap H$$

我们有

$$U = (U \cap G) \cup [U \cap \Gamma p(G)] \cup (U \cap H)$$

但是

$$U \cap \Gamma p(G) = a_0$$

由此

$$U = (U \cap G) \cup a_0 \cup (U \cap H)$$

我们考虑开集 $U \cap G$ 的边界.它包含在集合 U 和 G 的边界的和集中

$$\Gamma p(U \cap G) \subset \Gamma p(U) \cup \Gamma p(G)$$

但在集合 G 的边界点之中,只有点 a_0 可以加入到集合 $U \cap G$ 的边界内,而在 U 的边界点之中,只有属于集合 G 的点加入到 $U \cap G$ 的边界.因此

$$\Gamma p(U \cap G) \subset [\Gamma p(U) \cap G] \cup a_0$$

完全一样,我们也可以得到

$$\Gamma p(U \cap G) \subset [\Gamma p(U) \cap H] \cup a_0$$

因为集合 U 的边界是由 k 个点所构成,所以是有穷集,那么集合

$$G \cap \Gamma p(U) \text{ 和 } H \cap \Gamma p(U)$$

也是有穷的,这意思是说,集合

$$\Gamma p(U \cap G) \text{ 和 } \Gamma p(U \cap H)$$

更加是有穷的.

我们假定集合($U \cap G$)的边界是由 p 个点所构成,而集合 $U \cap H$ 的边界是由 q 个点所构成.

于是集合 $\Gamma p(U) \cap G$ 将包含不少于 $p-1$ 个点,而集合 $\Gamma p(U) \cap H$ 将包含不少于 $q-1$ 个点.

因为曲线 L 的每一点属于集合 G,H 或 $\Gamma p(G)$ 之一,而集合 U 和 G 的边界没有公共点,那么由此推出:集合 U 的每个边界点或者属于集合 G,或者属于集合 H.加入到集合 U 边界的点的数目等于 k;加入到集合 G 的集合 U 边界点的个数大于等于 $p-1$;属于集合 H 的集合 U 边界点的个数大于等于 $q-1$,所以

$$(p-1) + (q-1) \leqslant k < 2n - 2$$

由此
$$p + q < 2n$$
由这个不等式推出:数 p 或 q 至少有一个小于 n.

因此,我们看到:开集 $U \cap G$ 和 $U \cap H$ 之一,其边界是由少于 n 个点所构成,我们用 A_1 来表示开集 $U \cap G$ 和 $U \cap H$ 中其边界包含少于 n 个点的那个开集.因为
$$A_1 \subset U \subset \overline{U} \subset A_0$$
所以
$$\overline{A}_1 \subset A_0$$
更由于集合 U 的直径小于 $\frac{\delta_0}{2}$,那么集合 A_1 的直径就更加小于 $\frac{\delta_0}{2}$
$$\delta_1 < \frac{1}{2}\delta_0$$
最后,集合 A_1 的边界是由少于 n 个点所构成.

这样一来,我们就证明了:如果曲线在其每点具有支指数小于 $2n-2$,那么无论曲线 L 的开集 A_0 如何,总可以找到这样的一个开集 A_1,此开集以及其闭包皆含于 A_0 内,其直径小于 $\frac{1}{2}\delta_0$,而其边界包含少于 n 个点.

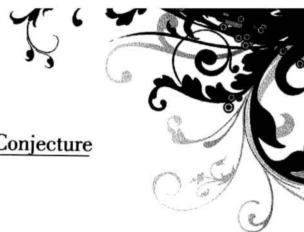
将此结果运用于集合 A_1,我们便找到一个集合 A_2,此集合以及其闭包皆包含在集合 A_1 中,其直径小于 $\frac{1}{4}\delta_0$,而其边界是由少许 n 个点所构成.继续作同样的推论,我们便得到一个开集降序列
$$A_1 \supset A_2 \supset \cdots \supset A_m \supset \cdots \qquad ②$$
其中每后一集合 A_{m+1} 及其闭包皆包含在前一集合 A_m 之中
$$\overline{A}_{m+1} \subset A_m$$
集合 A_m 的直径小于 $\frac{\delta_0}{2^m}$,而每个集合 A_m 的边界是由少于 n 个点所构成.由此推出:集合 A_m 的直径(即 $\overline{A}_m$ 的直径)随 m 增大而趋于零.因此集合
$$\overline{A}_1 \supset \overline{A}_2 \supset \cdots \supset \overline{A}_m \supset \cdots$$
形成一个降序列,此降序列具有唯一的公共点 x_0.

我们来证明曲线在点 x_0 具有支指数小于 n.无论正数 ε 如何,总可以找到自然数 m,使得
$$\frac{\delta_0}{2^m} < \varepsilon$$
序列 ② 的开集 A_m 具有直径小于 $\frac{\delta_0}{2^m} < \varepsilon$.集合 A_m 的边界包含少于 n 个点.

这样一来,曲线 L 在点 x_0 具有支指数小于 n.定理 7 就完全被证明了.

由所证明了的定理立刻推出:除了圆周以外,无论什么曲线也不能在它所有点具有同一有穷的支指数,因为如果在曲线 L 的所有点,支指数都等于 n,那





么在这曲线上就有一点,其支指数大于等于 $2n-2$,而当 $n \geqslant 3$ 时

$$2n - 2 > n$$

Π·C·乌利松曾证明,对于所有 n,都存在仅由支指数 n 和 $2n-2$ 的点所构成的曲线.特别,当 $n=3$,我们有 $2n-2=4$.仅由支指数 3 和 4 的点所构成的曲线的实例已在前面例 11 的按语中证明了.

乌利松也证明了有那样的曲线存在,即在其所有点皆具有无限的支指数,可列的支指数及连续统的支指数.希尔宾斯基覆盖可以作为在一切点具有连续统的支指数的曲线实例.

4.4 曲线的某些一般性质

在这一节中,我们将引入许多事实,这些事实深刻地揭示着曲线概念的本质,并且联系着曲线的一般定义和我们在具体曲线的考察中所作出的概念.在这里我们只限于定理的陈述,而不作定理的证明.

我们已经说过,乌利松所给出的曲线定义不仅包括平面曲线,而且也包括空间曲线.并且这定义可以逐字地搬到 n 维空间里去①.

但在这里要揭示出下面的一个重要事实,那就是,无论在什么样的空间里作了一曲线 C,则在普通三度空间里总有一曲线 C' 与曲线 C 同胚,就是说,用拓扑的观点来看,此二曲线是同一曲线.而且,不必在全三度空间的子集中间来寻找这些曲线,只消在三度空间的某个(其本身为曲线的)连续统的子集中间来寻找这些曲线.换句说话,在三度空间中有这样的"万有"曲线存在,它不仅包含三度空间的所有曲线的拓扑影像,而且包含 n 维空间的所有曲线的拓扑影像②.

因为太繁难,不能来证明这一件事实,我们只描述这种万有曲线的构造,此为奥地利数学家门格尔(K. Menger)所作.

把立方体分割为 27 个相等的立方体,挖掉它"中心立方体"及其毗连的 6 个立方体.结果便得到由 20 个一级闭立方体所构成的集合 V_1.对于剩下的每个一级立方体施行同样的手续,那就是说,把每个这样的立方体分为 27 个相等的立方体,且从每个一级立方体中挖去它中心立方体以及与之毗连的 6 个立方体.剩下的集合 V_2 是由 400 个二级立方体所构成.再对每个二级立方体作同样的处理,结果便得到闭集 V_3,此闭集 V_3 是由 $20^3 = 8\,000$(个)三级立方体所构成.此手续可以无限制地进行下去,至于任意自然数 n.

完成了所有这些手续以后,立方体 V_0 中所剩下的点集便是所求的"万有"曲线.作为连续统的下降无穷序列的公共部分,其本身为一连续统.

---

① 一般地说,此定义可以逐字地搬到其中界定了距离概念的任意集合,即任意度量空间里.
② 甚至是任意度量空间的所有曲线的拓扑影像.

证明这连续统是一样的,即证明这连续统确实是一曲线,那是非常困难的.

我们已经看到:希尔宾斯基覆盖起着平面曲线的“万有”曲线的作用.问题就发生了:为什么它不是所有一般空间曲线的“万有”曲线呢?原因就在于:在空间有这样的曲线存在,这曲线根本不可能拓扑地映象到任何平面曲线上.倘若我们考虑由四面体的六个棱边以及连接空间任一点和这四面体四顶点的四线段所构成的曲线,便得到这种曲线最简单的实例(图 19.21).

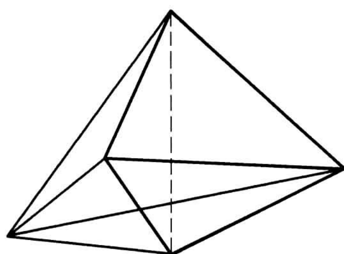


图 19.21

但如,四度空间的所有曲线都可以拓扑映象到三度空间的某曲线.当然并不能由此推出:一般三度空间的曲线具有那四度空间的指定曲线的所有性质.但是经拓扑映象而不损坏的那些性质是这二曲线的公共性质.

我们把曲线概念界定得很一般化,这就自然地引起一个问题:被我们定义做曲线的那些对象是否过于广泛了呢?现在援引曲线的两个性质以说明:称之为曲线的那些集合是多么适合寻常的曲线概念,这个概念首先与直线和圆周有所联系.

直线线段可以表示为有穷个任意小的部分线段之和,且曲线的任何点不属于两个以上的部分线段.任意曲线都具有类似的性质,甚至以此性质为曲线的特征.就是:欲使连续统 C 成为一曲线,其充分且必要的条件乃是:对于任意小的正数 ε,连续统 C 总可以表示为有穷个闭集之和,这些闭集的直径皆小于 ε,且连续统 C 的任何点不属于两个以上的这样的闭集.П·С·乌利松以更广泛的形式证明了这个定理,此定理特别断言:不成为曲线的连续统就不具有这种性质.比方说,如果取正方形并且把它分为相当小的闭集,那么总可以找到正方形的一点至少属于三个这样的闭集.然而总可以找到正方形的任意小分割,把正方形分为许多闭集,而使这正方形的任意点不属于三个以上的这些闭集(图 19.22).

重新回到曲线上来,我们看到在分割定理中闭集也可以不是连通的.欲使它们成为连续统,充分且必要的条件乃是:曲线是线段的连续影像.

使我们所引入的曲线概念和通常关于这对象的观念更加接近的另一性质是:曲线各点作任意小的移动,便可以使所有曲线变成折线.这一事实为苏联数



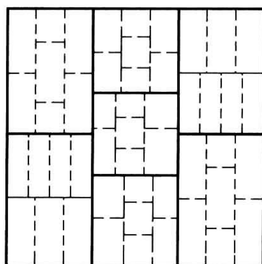


图 19.22

学家Π·C·亚历山大洛夫所揭示,在更一般的假设之下,此事实可以更确切地陈述如下:欲使列紧集 C 成为一曲线,必要且充分的条件是:对于所有任意小的正数 ε,可以把列紧集 C 映象到一个折线上,使这折线的每一点离集合 C 的相应点的距离小于 ε.

对于圆周来说,最简单的是在这圆周内作一正多边形,其边数充分大,且把每支圆弧(垂直地)投影到多边形的对应边上,便得到了这种映象.在一般情况下,也可以类似地来解决这个问题.其余前面的定理,把曲线分割为有穷个相当小的部分,使曲线的任何点不属于两个以上的部分.在这些部分的每个中取定一点.当两点所属部分有公共点时,便用直线线段连接之.这样一来,我们便得到一折线.把我们的曲线映象到这折线上,只要把曲线的分割取得相当小,便可以使得曲线的每一点借微小的移动而移到折线上的对应点.

5 关于维度的概念

曲线所有异于曲面、体乃至一般地异于任意连续统,主要是在于:曲线是一维的连续统(我们也说是一度).我们已经解释了这概念用于连续统的意义.

在这个补充里,我们要给出对于任意集合的维度的一般定义,此定义应该于Π·C·乌利松,我们将用许多例子来说明这个概念.为此,我们将假定集合 R 是在直线上,在平面上或在空间里(或甚至是在任意 n 度空间)①.我们不除去集合 R 和全空间一致的情形,因为下面所指引的维数的定义,不仅适用于指定空间的任何集合,而且也适用于全空间.维度的一般定义是按归纳法而引入的.

为了使它更明白,我们先研究维度.和维度 1 的个别情况,然后再界说一般的定义.我们说:集合 R 在点 x 具有维度 0,如果无论正数 ε 多么小,总可以找到一个包含着点 x 的开集 G,此开集 G 具有直径小于 ε,而其边界不含任一点(即空集).

① 在界说一般的定义和定理时,实质上我们将用到 R 为度量空间这一事实.

如果集合 R 在它的每一个点皆具有维度 0,那么我们就说集合 R 具有维度 0.

简单地界说了上述的定义,我们说集合 R 具有维度 0,如果其每一点包含在任意小(就其直径而言)的开集内,这开集的边界不包含任何一个点.回忆:开集不属于集合本身的那些极限点的总体叫做此开集之边界.我们也就可以说:如果集合 R 的每个点皆包含在任意小的开集中,而此开集同时也是闭集,那么集合 R 便具有维度 0.

我们援引几个例子,集合 R 是由一个点所构成,或是由有穷个点所构成,则集合 R 恒具有维度 0,因为在这个情况下,每个点既是开集,同时也是闭集(具有直径 0).

任意可列点集具有维度 0.事实上,设 x 是集合 R 的任一点,ε 是任意正数.我们用 r 来表示一个小于 $\frac{\varepsilon}{2}$ 的正数,且此正数异于从点 x 到集合 R 的各点之距离.因为集合 R 是可列的,而界于 0 和 $\frac{\varepsilon}{2}$ 之间的正数集合具有连续统的势,所以这样的正数是存在的.集合 R 中所有与点 x 的距离小于 r 的点的全体是一开集,它包含点 x 而其直径小于 ε,并且其边界不包含集合 R 的任一点,因为边界的所有点是在与点 x 的距离等于 r 的地方,而在 R 中没有这样的点.

由此推出:在任意度数的空间中,其所有坐标为有理数的一切点的集合将是 0 维度的,因为这样的集合恒为可列集.

特别,直线上一切有理点的集合具有维度 0.直线上所有无理点的集合也具有维度 0,因为我们无论取怎样的无理点 x 且无论选取什么正数 ε,总可以找到两个这样的有理数 a 和 b,使得下面的不等式成立

$$x-\frac{\varepsilon}{2}<a<x<b<x+\frac{\varepsilon}{2}$$

界于两点 a 和 b 之间的无理点的集合 G,是关于 R 的开集.

其边界不含任一点,因为不在 G 中的任何无理点皆不是 G 的极限点.集合 G 的直径小于 ε.

在关于 0 维度的集合的定理中,我们提一提下面的定理,今后我们需要它.

欲使一列紧集具有维度 0,必要且充分的条件是:此列紧集不包含任何连续统.

欲界定 1 维度集合的概念,我们需要刚才引进的 0 维度集合的概念,这是说明维度定义的归纳法的特点.

我们说集合 R 在它的一点 x 具有维度 1,如果下列两个性质同时成立.

(1) 无论正数 ε 怎样,总可以找到直径小于 ε 的开集 G,此开集 G 包含点 x,而其边界或者是空集,或者是 0 维度集.





(2) 存在这样的数 δ,在所有包含点 x 而直径小于 δ 的开集的边界上,至少有一个点.

如果只限于条件(1),那么 0 维度的集合以及 1 维度的集合都适合它;因此它表示出使集合 R 在点 x 具有维度小于或等于 1 的条件.条件(2) 是说集合 R 在点 x 的维度异于零.

我们说:集合 R 具有维度 1.如果集合 R 在其每一点的维度小于等于 1,即使在一点集合 R 的维度等于 1.

直线在其每一点具有维度 1.事实上,如果 x 是直线上的任意一点,而 ε 是任意正数,那么任意包含着点 x 而长度是 ε 的区间,都是合于 1 维度集合定义的条件(1) 的开集,这因为它的边界是由两个点(区间之端点) 所构成,故具有维度 0.另一方面,直线上的所有开集都是由一个或若干个区间所构成①.因此它的边界至少包含两个点,这意思是说,R 不是 0 维度集合.因此,1 维度集合定义的条件(2) 亦满足.所以,直线在它的每点具有维度 1.

用类似的方式,可以证明:圆周在它的每一点具有维度 1.

事实上,设 x 是圆周的任一点,ε 是任一正数.包含点 x 而弦长小于 ε 的圆周之弧段,对于圆周而言是开集,其直径小于 ε,而包含点 x,并且其边界由两点 —— 弧之端点所构成,因此具有维度 0.另一方面,因为圆周上的所有开集皆由一个或若干个弧段(无端点的) 所构成,所以圆周上任意开集的边界至少包含两点.由这些论证推出圆周在点 x 具有维度 1.

上面提到过:列紧集具有维度 0 的充分且必要的条件是,它不包含任何连续统,根据这个定理,现在我们可以说:在刚才所援引的定义的意义之下,曲线是具有维度 1 的连续统.

我们现在来谈维数概念的一般定义.

我们已经说过维数概念是按归纳法而引进来的,在这里归纳法宜于从数 -1 开始②.因而我们引入维度 -1 仅仅是为了叙述的方便,维度 -1 是加诸那种不含任何点的"集合".

在归纳法的定义的情况中总是这样做的,我们假定已经知道了维度 $0,1,2,\cdots,n-1$ 是加于什么样的集合,运用这些我们来确定维度 n 应该加于什么样的集合.

我们说:集合 R 在点 x 具有维度 n,倘若:

(1) 对于任意 ε,可以找到一个直径小于 ε 的开集,此开集之边界或是 0 维度的集合,或是 1 维度的集合,……,最后,或者是 $n-1$ 维度的集合.

(2) 有这样的 δ 存在,使得所有包含 x 点而直径小于 δ 的开集的边界,既不

---

① 构成直线上开集的区间集合,可能是有穷的,也可能是无穷的.

② 在以下的定义中,实际上并不是从 -1 维度开始的,而是从 0 维度开始的.

是 0 维度的集合，也不是 1 维度的集合，……，也不是 $n-2$ 维度的集合.

条件(1) 的意思是说：集合 R 在点 x 或具有维度 0，或具有维度 1，……，最后，或具有维度 n，因此，它表示出集合 R 在 x 点具有维度小于等于 n 的条件. 条件(2) 的意思是说：集合 R 在 x 点的维度不等于诸数 $0,1,\cdots,n-1$ 中的任意一个.

我们说：集合 R 具有维度 n，如果在其每个点它具有维度小于等于 n，而至少在一个点它的维度等于 n.

将这个定义用于1维度的集合. 条件(1) 在这里是说：对于任意 ε，可以找到一个包含点 x 而直径小于 ε 的开集，其边界具有维度 -1 或0；条件(2) 在这里的意思是：如果我们把 δ 取得充分小，那么所有包含点 x 而直径小于 δ 的开集的边界不应该具有维度 -1，即不应该是空集.

在维度 2 的情形，对于任意 ε，集合的每一点应该包含在直径小于 ε 的开集中，此开集之边界具有维度 -1，0 或 1；但是有这样的 δ 存在，使得所有包含点 x 而直径小于 δ 的开集的边界，即不是 -1 维度的集合，也不是 0 维度的集合.

比如，我们取平面. 因为它的每个点是直径任意小的圆的中心，所以可以说：平面的所有点皆包含在任意小的一个开集中，此开集的边界具有维度 1. 因此平面在它的每一个点具有维度小于等于 2.

欲证明平面在它的每一点的维度恰巧等于 2，只要证明：平面上的每个有界开集的边界既不是维度 -1 的集合，也不是 0 维度的集合.

倘若平面上的开集的边界是空集，这意思是说：这开集同时是闭集，于是它的余集也将是开集同时是闭集，因此，全平面便分割为两个无公共点的闭集. 但这是不可能的，因为平面是连通的集合.

平面上的有界开集的边界不可能是 0 维度的集合，这也是正确的.

事实上，平面上所有有界开集的边界是有界闭集，故具有列紧性. 如果假定它具有维度 0，那么根据上面提到的关于零维度集合的定理，它不可能包含任何连续统.

但在以前我们说过，倘若平面上的列紧集不包含任何连续统，那么它就不分割平面，但平面上一切开集 G 的边界分割平面为两个无公共点的开集 G 和 $R\setminus\overline{G}$

$$R\setminus \Gamma p(G) = G\cup(R\setminus\overline{G})$$
$$G\cup(R\setminus\overline{G}) = 0$$

因此平面上有界开集的边界不可能是 0 维度的集合. 如果我们想要证明空间在它的每一点具有维度 3，那问题是极其复杂的. 于是，在三维空间的情形和 n 维空间的情形之间没有本质的区别. 因此我们立刻转向 n 维的情形.

为了验明 n 度空间在乌利松的意义下具有维度 n 而必须遵循的方法，在这





里不可能描述,我们只列举出维度理论中的一些定理,从这些定理可以得到这个基本结果,顺便认识一下拓扑学的这个分支中的最重要的事实.

下列二定理很简单地就可以得到证明:

定理 1　如果集合 R 的维度等于 n,那么集合 R 的一切子集 R' 的维度可能小于 n 或者等于 n.

定理 2　两个彼此同胚的集合具有相同的维度.

特别,如果集合 R 在 x 点具有维度 n,在指定的同胚映象之下,和集合 R 同胚的集合 S 在对应于点 x 的点 y,也具有维度 n.

定理 1 和定理 2 可以用归纳法证明.运用这两个定理并且按归纳法进行推论,不难证明下面的定理 3 成立:

定理 3　n 度空间在它的每一点具有维度小于等于 n(在乌利松的意义之下).

困难在于证明 n 度空间的维度确切地等于 n.有许多不同的方法来证明这个事实.我们拟定其中的一个方法如下:

首先证明这样的一个一般定理:

定理 4　如果列紧集 R 在它的每一点具有维度小于等于 n,那么对于任意 ε,它总可以用闭集的一个有穷组来遮盖,这个组中每个闭集具有直径小于 ε,而集合 R 中没有一点属于这个组的多于 $n+1$ 个集合.

我们不证明这个定理了,只援引一些例子来说明这个定理.

首先我们取直线线段为集合 R,它在其每一点具有维度 1.将它分成相等的子线段,使每个子线段的长度小于指定的正数 ε;于是线段的每一点或是小线段的内点,或是两个这种线段的公共端点.

现在我们取正方形,和全平面一样,它在其每一点具有维度 2.对于任意 ε,总可以把它分成许多长方形,这些长方形的边平行于正方形的边,而其对角线的长度小于 ε.于是,正方形没有一点关于三个以上的这种长方形(图19.22).

最后,如果我们取正方体为集合 R,那么可以把它分成长方体,其边平行于正方体的各边,而对角线的长度小于 ε,并使正方体没有一点属于四个以上的长方体.

定理 4 的逆定理成立:

定理 5　如果对于任意 ε,列紧集 R 可以分割为直径小于 ε 的开集,而使得集合 R 中没有一点属于多于 $n+1$ 个这种闭集,那么整个集合 R 具有维度小于等于 n.

比较正定理和逆定理的论断,我们便得到下面的结论:

定理 6　欲使列紧集 R 具有维度 n,必要且充分的条件是:对于任意 ε,集合 R 可用直径小于 ε 的闭集组来遮盖,而使得集合 R 没有一点属于这个组中

$n+1$ 个以上的集合;但是有这样的 δ 存在,使得如果组的所有集合皆具有直径小于 δ, 那么这些集合中有某 $n+1$ 个具有公共点.

对于整个维度理论而言,这个性质具有很重要的价值,因而人们常常把它取作列紧集维度的定义.如果对于任意 ε,列紧集 R 可以用直径小于 ε 的闭集组来遮盖,而使得集合 R 没有一点属于这个组中 $n+1$ 个以上的集合,但如果把 ε 取得充分小,则在 R 中恒有一点属于这个组的 $n+1$ 个集合,就说列紧集 R 具有维度 n.

在确定了维度 n 的集合的一般性质以后,便证明一切 n 维单纯形(симплекс)①皆具有这个性质.十分容易证明:对于任意 ε,可以把单纯形分成有穷个直径小于 ε 的闭集,而使得单纯形没有一点属于此分割中 $n+1$ 个以上的集合.为此,只要把单纯形包含在一个 n 维正方体中,并且将这正方体分割为许多长方体,使这些长方体对角线的长度小于 ε,而这些长方体中没有 $n+2$ 个是具有公共点的,并且取单纯形和每个长方体的公共部分作为单纯形的分割的元素.

最困难的部分是证明下面的一个事实.就是有这样的 δ 存在,如果单纯形所分割成的那些闭集的直径小于 δ,那么在这些闭集中就可以找到具有公共点的 $n+1$ 个闭集.

因为单纯形是列紧集,所以参照一般的定理 6 及单纯形的这个性质,便推出:n 维单纯形至少在其一点具有维度 n(在乌利松的意义下).命这一点是 x;那么不仅单纯形,而且全空间在点 x 也具有维度 n.事实上,我们已经看到,n 度空间在其每一点的维度小于等于 n(定理 3).

因为单纯形在 x 点的维度大于等于 n,所以空间在点 x 的维度不可以小于 n.由此断定:空间在点 x 的维度恰好等于 n.

最后,空间的所有点 x 可以借同胚映象而转移到另一点 y,在这种意义之下,n 度空间是均匀的.这只要考虑把点 x 移到点 y 的平行移动就够了.然而空间在点 x 具有维度 n,所以按照定理 2,它在点 y 也将具有维度 n.因为点 y 是任意选取的,所以由此推出:n 度空间在其每一点皆具有维度 n(在乌利松的意义之下).

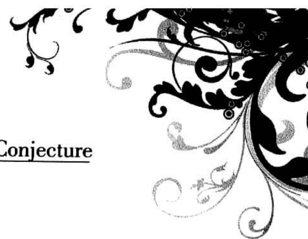
我们已经考虑了一个进行证明下列事实的方法,这事实是 n 度空间在其每一点皆具有维度 n(在乌利松意义之下).还有许多证明这个定理的其他方法,但所有的这些方法,如我们所说,皆具有困难的成分.试问,如果只是为了证明空间具有维度 3(而这情形丝毫也不比 n 度空间的一般情况更简单一些),值不

---

① 点是零维单纯形,线段是一维单纯形,三角形是二维单纯形,四面体是三维单纯形.n 维单纯形可以定义为 n 度空间中其坐标适合下列不等式组的点的集合

$$x_1 \geqslant 0, x_2 \geqslant 0, \cdots, x_n \geqslant 0$$
$$x_1 + x_2 + \cdots + x_n \leqslant 1$$





值得引进这样的维度的定义,而必须建立起一整套理论呢!

可以用下面的几个事实来说明这个定义是恰当的.首先,借助于这个定义,可以按其维度的观点而来描述 n 度空间的任意集合的特征(在曲线的例子上我们已经看到了这一点);其次,只要运用这个维度定义,我们就可以证明:具有不同维数的两个空间不可能互相单值互相连续地由一个映象到另一个上①,至少这一点是重要的.参照以下的两个事实,很简单地就推出这个命题.在乌利松的意义之下,n 度空间具有维度 n;互相同胚的两个集合具有同一个维度.特别,如果希望空间的点连续地依赖于它的坐标,而坐标又连续地依赖于点,那么空间的每一点应该用三个坐标给出.欲证明这一件事实,我们所引进的定义是必要的.

① 这个命题叫做维数的不变性定理,它对整个数学而言,具有很重要的意义.

直觉的讨论[①]

第二十章

1 拓扑学的主要问题

拓扑学(Topologie)所讨论的对象,是几何图形的,经过拓扑变换(topologische abbildung)—— 正逆两方面都单值(Eindeutig)而又都继续的变换 —— 而不改变的性质.我们暂时先把几何图形就当做是三维空间(或更高维的空间)中的点集.在空间的一笛卡儿坐标系中,连续函数所表出的变换叫做连续变换.这些变换函数 —— 表出变换的函数 —— 只要在这图形的点集上有定义(不必在全个的空间中都有定义).图形的性质,经过拓扑变换而不改变的,叫做拓扑性质.两个图形间若是有拓扑对应,那就是说,若有一拓扑变换存在,把两个图形中的一个换成另一个,这两个图形就叫做同胚的(Homöomorph)图形.

例如半个球面与圆域(Kreisscheibe)同胚,因为直角投影(Orthogonalprojektion)就是把半个球面换成(图 20.1 中用斜线表示的)圆域的一个拓扑变换.更普遍地说,若是一曲面能弯扭成另一曲面,它们就同胚.例如球面、立方形,与椭圆面同胚;平环(Kreisring)与有限高的圆柱面也同胚.

① 本文节选自 H·赛费特和 W·特雷法尔合著的 *Lehrbuch der Topologie*.
赛费特(Seifert. Herbert.K.I.1907),德国数学家,海德堡大学教授.主要贡献在拓扑学方面.
W·特雷法尔(Threlfall. William R.1888—1944)德国数学家,生于德累斯顿(Dresden).主要贡献在拓扑学和变介法方面.



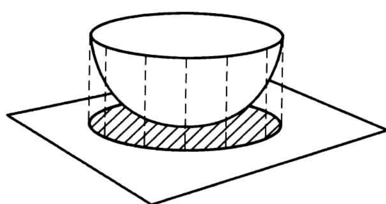


图 20.1

同胚图形的例子多而易见.但是也有图形,例如欧几里得平面与刺破了一个点的球面(减去了一个点的球面),不是一见就可以断定它们同胚的.这两个图形间的画形投影(stereographische projektion)就是一个拓扑变换,而且它们也与圆域的内域同胚.

要证明两个点集不同胚,通常比较困难.点与线段显然不同胚,因为这两个点集不能成一一对应.我们也容易看出线段与圆域不同胚.设 A,B,C 是圆域中的任意三点.从点 A 出发,能绵续地在圆域中走到点 B,而不通过点 C.圆域的这个性质是拓扑性质,线段却没有这个性质:若要从线段的一端绵续地走到另一端,就必须通过线段的中点.若更进一步比较圆域与球体,我们就不能如此简单地得着结论.圆域中有分开圆域的闭曲线;若是我们要想利用圆域的这个特征,我们就必须证明一条闭曲线不能分开球体.但是为什么一条闭曲线不能装满一个分开球体的曲面呢?我们要证明圆域、球体,与更高维的相当的图形不同胚,已经与证明线段与圆域不同胚的情形不同,不能有那么简易的方法了.

同胚的概念在拓扑学中的地位,与全合的概念(Kongruenzbegriff)在初等几何学中的地位一样.在初等几何学中,两个全合的图形本质上无区别;同样的,在拓扑学中,两个同胚的或拓扑对应的图形本质上也无区别.但是我们应当注意下述不同之处:两个图形若是全合,全个的空间就有一个刚体运动,使其中的一个图形移到另一个;两个图形若是同胚,全个的空间却不必有一个拓扑变换,使其中的一个图形换成另一个.

例如全个的空间的任一拓扑变换都不能(任一变状(Deformation)当然更不能)把圆周(Kreislinie)换成一个扭成结的(Verknotet)曲线例如三叉扭结(Kleeblattschlinge)(图 20.2).但是圆周与扭结(Knoten)同胚,因为它们的点间有一个正逆两方面都单值而又都绵续的变换,而且同胚这种关系只依赖于这两个点集,与那包含它们的空间无干.同样的,全个的空间没有一个拓扑的自身变换(Selbstabbildung),把一个扭转了 2π 的倍数的弧度的环带(geschlossenes band)(图 20.3)换成一个未扭转的环带.但是把这两个环带割开成两个全合的长方条,使条上相当的点成对应,就可知这两个环带同胚.

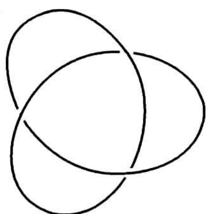


图 20.2

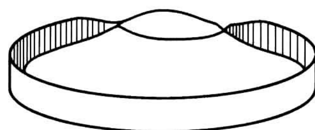


图 20.3

在拓扑学中,扭结与圆周,或如此扭转了的环带与未扭转的环带,都是相同的图形.只有把它安置在三维空间中的时候,它们才有区别;而且若是把那包含它的三维空间又当做是四维空间的子空间,因而可以引用四维空间的变状,这种区别可又消失了.这是因为经过四维空间的一个变状,圆周可以换成扭结,而且在变状的历程中曲线还不自相穿割(Selbstdurchdringung);就如同经过三维空间的一个变状,圆周可换成椭圆.

由此看来,此后我们要把一个图形的拓扑性质分成两种:一种是"内在的",是经过这个图形的所有的拓扑变换都不改变的性质;另一种是依赖于包含这图形的空间的,是经过这全个空间的所有拓扑变换都不改变的性质.

我们再举一例来说明这种区别.设空间中有一圆周与一直线,同在一平面上,但无共点.圆周绕直线旋转,产生一个环面(Ringfläche 或 Torus).通过环面的任一点 O 有一母圆 a;我们把它叫做环面的经圆(Meridiankreis).在旋转时,点 O 也产生一圆 b;我们把它叫做纬圆(Breitenkreis)(图 20.4).我们能不用空间的拓扑的自身变换,表出经纬圆的区别如下:经圆能在环体(Vollring,环面所包的立体)中缩成(Zusammenziehen)一点,而纬圆不能,所以全个的空间的任一拓扑的自身变换,都不能把由环面与一个经圆所构成的图形,换成由环面与一个纬圆所构成的图形,但是经纬圆间的这区别并非环面的内在性质.空间虽然没有一个变状,把环面换成它自己,使经纬圆的地位交换;环面却有如此的一个拓扑变换.要想求得如此的一个拓扑变换,我们设想这环面是用有弹性的薄膜做成的.把它沿着 a 与 b 割开(图 20.5),弯扭成一个正方形;再把正方形沿着对角线折叠起来,使正方形上原来的两点,折叠成一点的,成对应.这对应就是正方形的一个拓扑的自身变换,使 a 与 b 交换,而且相当于环面的一个同样的自身变换.—— 将来在下节中,我们要提出曲面的能定向性(Orientierbarkeit) 与双侧性(Zweiseitigkeit).这个性质供给我们另一个特征,表明内在的拓扑性质与浸没的拓扑性质(Einbettungseigenschaft) 的区别.

拓扑性质间的这种区别,与微分几何学中的度量(Metrisch) 性质间的区别相同.微分几何学中的度量性质的一种是内在的,不依赖于曲面在空间中的位



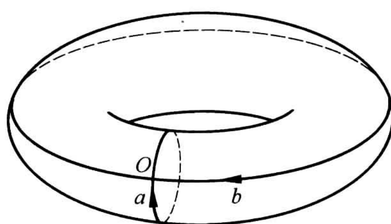


图 20.4

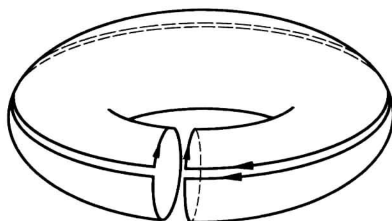


图 20.5

置,由第一个度量的基本式(Grundform) 断定;另一种是由曲面与空间所组成的图形的度量性质,由第二个基本式断定.

拓扑学的主要问题,就是要判断给定的两个图形是否同胚,而且在可能的时候,列举所有不同胚的图形.我们起始时就把图形看做是欧几里得空间中的点集.虽然关于欧几里得空间中的任意子集已经有很深博的理论,为避免引起不愿意引起的集合论中的困难起见,我们将不再采用如此广义的图形概念了.我们将要采用的,只是 L.E.J.Brouwer 所创始的复合形(Komplex) 的概念.复合形在本书讨论的历程中再加以限制,使成为流形(Mannigfaltigkeit).这种概念不过于宽泛,因而可以避免所不愿引起的集合论中的困难;也不过于狭窄,因而可以包含差不多所有的有趣的图形.所以我们现在要研究的拓扑学,不是集合论的,只是复合形与流形的拓扑学.

复合形的特征,使它有别于任意点集的,是它的能剖分为三边形的性质(Triangulierbarkeit):复合形是由有限个或可数的无穷多的,不必平直的(Geradlinig),线段、三边形、四面体与高维的相当的图形连接而成的.这种三边形等的连接,并不限于在同一空间中,而且有时候还不要包含它们的空间.因为这特征,大部分所谓病态的点集都被划分在我们讨论的范围之外,而且我们的讨论才与几何的直觉更为接近.因此也有人把复合形的拓扑学叫做弹性橡皮的拓扑学.复合形的例子:所有的 Riemann 曲面,任何维的欧几里得空间,投影平面与投影空间,所有的欧几里得与非欧几里得的空间型(Raumform),度量的运动群的间断域(Diskontinuitätsbereich),与力学系统(mechanischer System) 的位置空间(Lagenraum) 与相空间(Phasenraum).

我们要想向着我们的目的 —— 主要的问题的解决 —— 进展,我们必须探求复合形的拓扑不变的,能计算的,而且能作为分类的标志的性质.复合形的同调群(Homologiegruppe) 与基本群(Fundamental-gruppe) 是这种性质中最重要的,而且在我们讨论中占中心地位.

同时我们要看看,不用这些不变性的帮助,我们能进展到如何程度.所以我们不再从事预备,立即直觉地讨论主要问题的一部分,看看有些什么不同胚的闭曲面.

2 闭曲面

在前一节中,我们已经说过,环面沿着经纬圆切开之后,变成一个正方形.反之,若是叠合一个正方形的每两条相对的边,我们又重得着环面.所以一个正方形的每两条相对的边当做是同一条线段的时候,这正方形就叫着环面的庞加莱基本多边形.至少对于内在的拓扑性质说,这基本多边形就可以完全代表环面.至于环面的面积,空间中的位置等特别的度量性质与浸没性质,基本多边形却不能表出;不过这些性质与曲面拓扑学是毫无关系的.所以从曲面的拓扑学的立场说,所有的环面,由叠合基本多边形的对应边而得着的,都等价(gleich wertig),例如旋转环面与扭成结的皮管子的表面就没有区别.其实任一曲面都能够切开成一个或数个多边形.我们要直接利用这事实,来下闭曲面的定义:一个闭曲面是由一对一对的连接有限个多边形的边而成的.这定义才使闭曲面的图形脱离了那包含它的空间,才使它成为不依赖于那包含它的空间而存在的二维流形 —— 二维流形的概念,将来有正确的定义.

根据环面切开成正方形的方法,我们能用若干个多边形,设想它们的边成对的对应,表出一个闭曲面,因而得着一组无限多个闭曲面.为达到这目的起见,我们先在环面上穿一个大致是圆形的洞(Loch).假使洞边缘(Lochrand) l 通过 O 这个点.经过变状之后,我们得着一个环柄(Henkel)(图 20.6).我们也能先在环面切开成的正方形中表出这个圆洞(图 20.7),然后把洞边缘 l 在点 O 处切断.这就等于把环柄沿着曲线 a 与 b 切开成五边形(图 20.8).这五边形的 l 这边不与别个边成对应,其余的都间隔的,成对的对应.



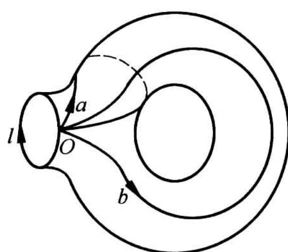


图 20.6

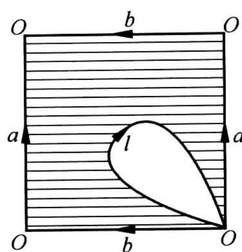


图 20.7

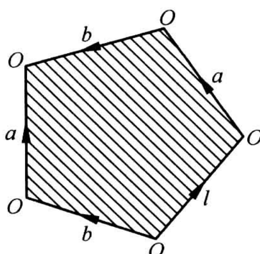


图 20.8

若是取两个切开成五边形的环柄，连接它们的洞边缘(图 20.9)，然后消去公共洞边缘 l，结果是一个八边形，它的边成对的对应(图 20.10). 叠合对应边，因而叠合 8 个顶点成一点，这八边形就包成一个双环面(Doppelringfläche). 与从前讨论环面时一样，我们要在图中表出成对的边的对应关系：每对对应边我们用同一个字母表示；而且用箭头附在边上，表明在叠合对应边时，箭头指着的方向应该符合. 我们只要任意地选定多边形的边缘的一个方向，然后按照每一边上的箭头是否指着这个方向，规定这边附带一个指数 $+1$($+1$ 这指数常省去)或 -1. 边的对应关系，以及这闭曲面，都能用一个式子表出. 例如一个双环面，若用图 20.10 中所表出的边的对应关系，而且选定那八边形的边缘的一个适当方向，它就可以写成下式

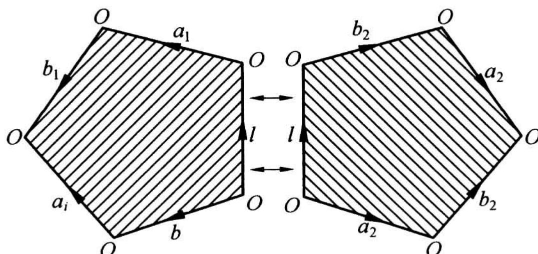


图 20.9

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

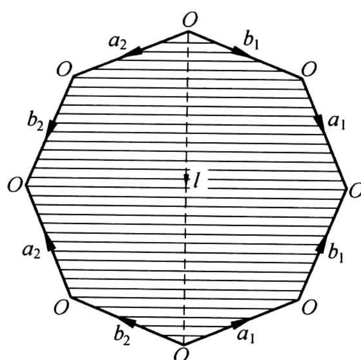


图 20.10

$$a_1 b_1 a_1^{-1} b_1^{-1} a_2 b_2 a_2^{-1} b_2^{-1}$$

再在双环面上穿一个圆洞,使洞边缘通过点 O,沿着这个洞边缘,还可以连接上另一个环柄.消去洞边缘之后,我们得着一个十二边形.如同双环面可以变状(变状不毁坏拓扑性质)成一个安装上了两个环柄的球面,这十二边形包成一个安装上了三个环柄的球面.我们能如是得着的曲面,是能切开成 $4h$ 边形的,安装上了 h 个环柄的球面.因为能任意变状,这曲面也可以说是安装上了 $h-1$ 个环柄的环面.下式

$$a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_h b_h a_h^{-1} b_h^{-1} \tag{h}$$

确定它的基本多边形 ——$4h$ 边形 —— 的边的对应关系.若是沿着多边形的一条对角线 l 把 $a_i b_i a_i^{-1} b_i^{-1}$ 四条边切下,我们就得着一个环柄切开成的五边形.所以每一段 $a_i b_i a_i^{-1} b_i^{-1}$ 相当于一个环柄.

球面也有一个确定的基本多边形;沿着球面上连接 O 与 P 两点的弧 a 把球面割开.结果是一个二边形,它的边缘圆(Randkreis)是

$$aa^{-1} \tag{0}$$

这基本多边形上有两个不相同的,即不相叠合的顶点 O 与 P(图 20.11),别个基本多边形都不如此.叠合这二边形的两边,使圆域再变成球面,就如同把张开的钱包在 O 与 P 两个铰链间的口子合拢起来,使它变成球形袋.

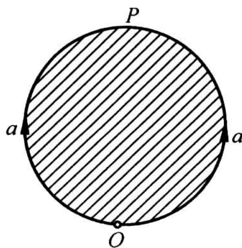


图 20.11

安装上了 h 个环柄($h = 0,1,2,\cdots$)的球面只是所有的不同胚的闭曲面的半数. 从球柄着手, 我们得着这半数的闭曲面; 同样的, 从 Möbius (带 Möbiusband) 着手,我们才可以得着其余的半数. Möbius 带是一个扭转了 π 弧度的环带. 图 20.12 所表示的 Möbius 带,可说明如下:设 O,P,M 分别是线段 c 的起点,终点与中点;设另有一轴与 c 无共点. 使 c 绕轴旋转,适合下列二条件:第一,c 与轴同在一平面上;第二,当这平面旋转一周的时候,c 在旋转平面中绕 M 旋转的弧度恰等于 π.

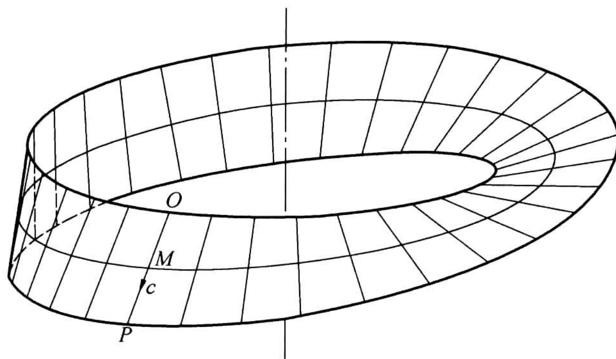


图 20.12

若是把 Möbius 带沿着线段 c 切开,结果是一个矩形. 它的边缘圆是

$$cr'cr''^{-1}$$

而且它们有两个不相同(不相叠合) 的顶点 O 与 P(图 20.13). r' 与 r'' 二自由边连成一闭曲线或拓扑圆,是 Möbius 带的边缘. 其余两边 c 与 c 的指数并非相反,与从前讨论的基本多边形的边不同;两边 c 与 c 的指数相同;表示它们的箭头指着边缘的同一方向. 两边的这种对应叫做反对应或第二种对应,而从前的那种叫做正对应或第一种对应.

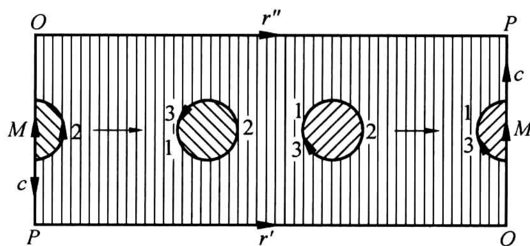


图 20.13

环柄与 Möbius 带的边缘都是一个拓扑圆. 但是它们有一个重要的区别:环柄在欧几里得空间中是双侧的(Zweiseitig),Möbius 带却是单侧的(Einseitig). 双

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

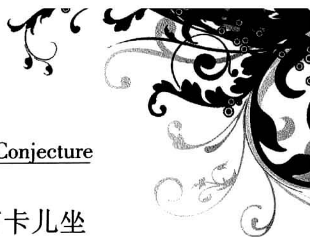
侧与单侧的意义如下:设 Möbius 带上有一只小虫沿着带的中线爬行,它就可以不越过带的边缘爬到它的起点的反面.环柄的边缘隔开环柄的两面;我们能把一面涂成白色,另一面涂成黑色,而且使黑白两色只在沿边缘处衔接.Möbius 带却非如此,它只有一侧.

这单侧性还不定是曲面的内在性质,所以还不能根据单侧性,以为一个扭转了 π 弧度的环带不能与一个没有扭转的环带同胚.这么抗议是很理的.不过根据另一内在的性质,这两种环带确实不同胚.这性质可说明如下.设想不是一只小虫在曲面上爬行,却是一个带有箭头的,带有 1,2,3 三个点的,小圆在曲面中称动.若是我们的曲面是 Möbius 带,而且毫无厚度,我们就可以移动小圆,使它与原来的小圆叠合,但箭头相反,1,2,3 三点与 3,2,1 三点分别相合(图 20.13).若是在一曲面中能如此移动一个小圆,我们就说这曲面不能定向(Nichtorientierbar);否则说这曲面能定向(Orientierbar).在能定向的曲面中,一个带有箭头的小圆,不但表明一个点的邻域的定向(Orientierung),而且这一点处的定向可以确定地移到每一点处去,能定向性是曲面的一个内在性质.反之,只有曲面浸没在三维空间中的时候,曲面的双侧性才有定义;所以双侧性是依赖于浸没的情形的,不容与能定向性相混乱.其实将来还要证明,能定向的曲面虽然永不会是一个浸没在欧几里得空间中的单侧曲面,但还可以是一个浸没的别个三维空间中的单侧曲面.

若是环柄上的圆洞用一个圆域盖上,或者换个说法.把环柄上的洞边缘连接到一个穿了洞的球面的边缘上去,我们就从环柄得着一个环面.同样的,也把 Möbius 带的边缘用一个圆域盖上,那就是说,把 Möbius 带的边缘连接到一个穿了洞的球面的边缘上去,我们就从有边缘的 Möbius 带得着一个闭曲面.在三维空间中,只有听任曲面自相穿割的时候,才能如此连接 Möbius 带与穿了洞的球面,否则,是断然不能实现的.不过我们能证明,在四维空间中,即使不听任曲面自相穿割,也能使这种连接实现.所以闭 Möbius 带是四维空间中的一个曲面.但是对于这曲面的内在性质说,它能否浸没在空间中完全无关重要;我们只要在 Möbius 带的边缘与穿了洞的球面的边缘间建立了一个拓扑对应,把每对对应点看做是同一点,闭 Möbius 带的内在性质就能完全断定了,而且在每一点的邻近处,闭 Möbius 带与平面的面片(Flächenstück)并无区别.至于如此界说的曲面能否全个地在空间中浸没,与能否在空间中把两个曲面沿着边缘连接起来等问题,我们并不加以讨论.

如此封闭了的 Möbius 带就是投影平面(projektive ebene).算学中所讨论的闭曲面中,投影平面的重要,仅次于球面.投影几何学中介绍投影平面的时候,并不说这是闭 Möbius 带,却说它是用一条假(Uneigentlich) 直线封闭了的欧几里得平面.根据后一种说法,在一个投影坐标系中,投影平面的点与三个实数的





比 $x_1 : x_2 : x_3$(0 : 0 : 0 除外)成一一对应.若是 x_1, x_2, x_3 当做齐次的笛卡儿坐标①,$x = \frac{x_1}{x_3}, y = \frac{x_2}{x_3}$ 就是欧几里得平面中通常的笛卡儿坐标,方程式 $x_3 = 0$ 就代表那封闭欧几里得平面而成投影平面的假(无穷远)直线.若是把 x_1, x_2, x_3 当做三维空间中的笛卡儿坐标,投影点就与通过原点的直线成一一对应.

我们现在要证明这线集与那组成闭 Möbius 带的点集,在拓扑性质方面,完全没有区别.取一个用原点做中心,半径等于1的球面.这锥线集(Büschel)中的每一直线交球面于两个径点.这就把投影平面的点拓扑地换成了单位球面上的径点耦.所以只要把球面上的每两个径点当做是同一个点,就得着投影平面.投影平面因此可以表出如下:用下半个球面的点做投影点的代表,而且设想在越过下半个球面的边界 —— 赤道圆 —— 的时候,能从这圆上的一点跳到它的径点.设平面 E 是球面的南极 S 处的切面.现把下半个球面直角的投影到平面 E(图 20.14),把它换成单位圆域.所以单位圆域的边缘上每两个径点叠合之后,这单位圆域变成的闭域就代表投影平面.

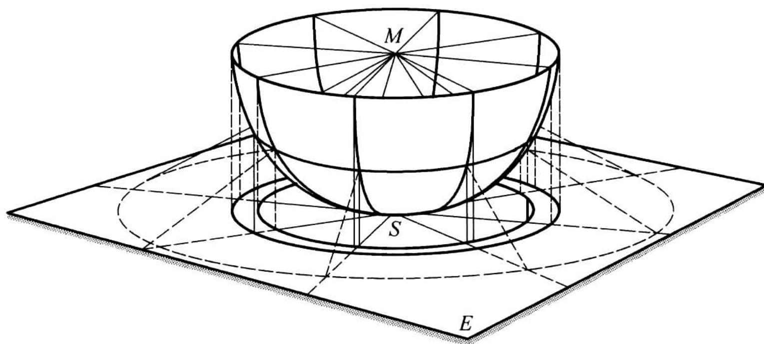


图 20.14

根据这个变换,我们可以得着两个结果.第一,投影平面与闭 Möbius 带相同.设 r' 与 r'' 是圆域中与圆心等距的二平行线段.沿着这二线段把圆域切开成下列三部分,两个弓形与其间的一个面片(图 20.15).叠合单位圆域的边缘上的径点,恰使这面片变成一条 Möbius 带,而且使这两个弓形沿着 b 边连接而成一个圆域.这 Möbius 带与这圆域的边缘都是 $r'r''$;所以后者就是封闭 Möbius 带的圆域(图 20.16).

闭 Möbius 带既然是投影平面的拓扑的代表,Möbius 带也可以从当做一个

① G. Kowalewski, Analytische Geometrie(Leipzig 1923), § 14, p. 32.此处所需要的投影几何学中的基本事实,在下列各书中都有适当的叙述,读者可以参考:F. Klein, Nichteuklidische Geometrie(Berlin 1928),章 Ⅰ;L. Bieberbach, Projektive Geometrie(Leipzig 1931);Hilbert Cohn-Vossen, Anschauliche Geometrie(Berlin 1982);H. Weyl, Die Idee der Riemannschen Fläche(Leipzig 1923).

穿了洞的投影平面.这是 Möbius 带的一个新表示法.洞不管在何处,结果所得的有边缘的曲面在拓扑学中都等价.所以若是叠合径点的一个圆域表示投影平面,我们就可以在圆域的中心穿一个圆洞,得着如图 20.17 中所表示的 Möbius 带.图 20.17 中的平环的外边缘的径点叠合,内边缘是 Möbius 带的边缘.若是再把这平环沿着虚线切开,然后叠合外边缘上的两个半圆,就重回到 Möbius 带的原来表示法了.

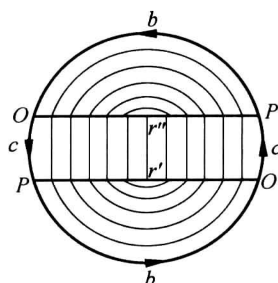


图 20.15

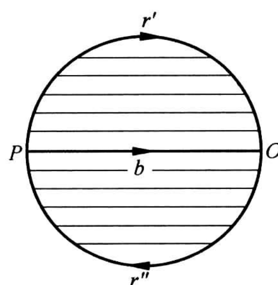


图 20.16

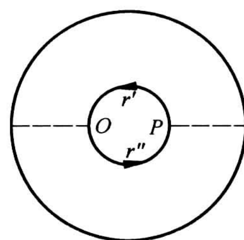


图 20.17

第二,投影平面换成圆域的交换立刻使我们求得投影平面的基本多边形.我们只要把单位圆域看做是一个二边形,两条边都是 a 的二边形.边缘上径点的叠合可用下式

$$aa$$

表出.a 边的起点与终点叠合而成投影平面上的点 O(图 20.18),而 a 就是一条投影直线的象(Bild) 线.把投影平面沿着这条投影直线割开,我们就得着投影平面的基本多边形.

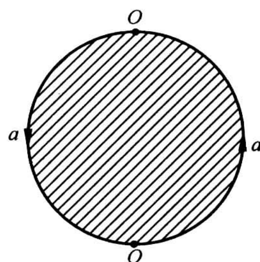


图 20.18

因此投影平面也是能用一个基本多边形代表的这一组闭曲面中的一个.它与这一组中的环面不同;环面是欧几里得空间中的闭曲面;但是,若是不准自相穿割,投影平面 —— 其实所有的不能定向的闭曲面 —— 却不能浸没在欧几里得空间中.

因为 Möbius 带,穿了洞的投影平面,与外(或内) 边缘圆上径点叠合之后的平环,这三个曲面同胚,所以在任一穿了洞的闭曲面上,沿着这个洞边缘连接上一条 Möbius 带或连接上一个穿了洞的投影平面,或叠合洞边缘上每对的径点,我们都得着同一个新曲面,这安装在圆形洞上的 Möbius 带有时候也叫做交叉帽(Kreuzhaube).



现在只要在球面上安装上若干 Möbius 带,我们剩下而还未讨论的闭曲面就能全体得着了.安装 Möbius 带的历程与安装环柄的一样:在投影平面中穿一个洞(图 20.18),使洞边缘 l 通过点 O.若是在点 O 处把 l 切断,从投影平面的基本多边形就得一个三边形,它们的边缘圆是 aal.设想有两个三边形,它们的边缘分别是 a_1a_1l 与 $a_2a_2l^{-1}$.把它们沿着洞边缘连接起来,再把连接的线痕消去,我们就得一个四边形,它的边缘圆是

$$a_1a_1a_2a_2$$

(图 20.19).这就是安装上了两条 Möbius 带的球面的基本多边形.

这曲面也叫做单侧的环管或不能定向的球面.若是不沿着连接的线痕 l,却沿着另一条对角线 m,把基本多边形切开成两个三角形,再沿着 a_2 边把它们连接起来,我们就得一个新的四边形(图 20.20),它的边缘圆是 $a_1ma_1^{-1}m$.这四边形所包成的当然是同一个曲面.我们先叠合 a_1 这条边,使这四边形变成圆柱面(图 20.21).要得着我们的曲面,圆柱面的两个界圆 m 上的点必须成对地叠合.但是我们不能叠合每一条母线上的两个端点,因为如此叠合而成的是能定向的环面.设 x 是圆柱面的轴的平分垂线.我们必须叠合两个界圆上的,对于 x 对称的每两个点,才得着我们的不能定向的环面.

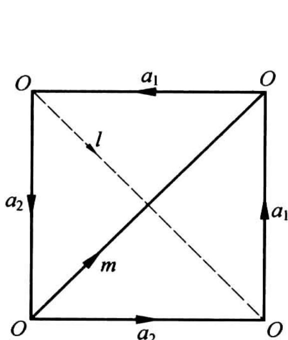


图 20.19

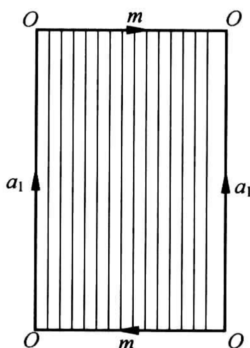


图 20.20

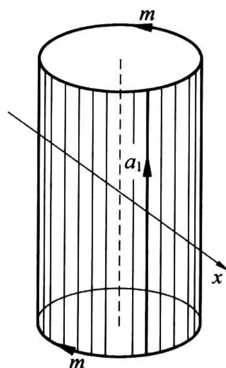


图 20.21

若是听任圆柱面自相穿割,这两个界圆的这种叠合也可以在欧几里得空间中实现(图 20.22).如此得着的单侧的环管从顶到底的切开,结果是两条 Möbius 带.图 20.23 表出其中的一个.

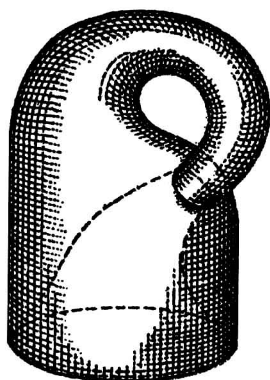


图 20.22

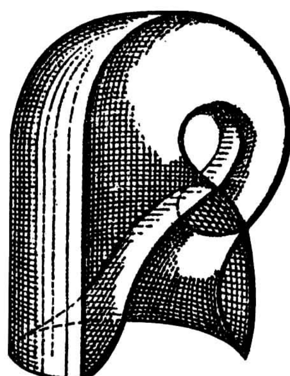


图 20.23

球面上安装 k 个 Möbius 带所成的闭曲面，可以展开成一个基本多边形，它的边缘圆是

$$a_1 a_1 a_2 a_2 \cdots a_k a_k \tag{k}$$

我们现在既不知道，我们如此得着的这些闭曲面是否是所有的闭曲面，也还没有证明这些闭曲面中没有两个同胚.第一个疑问很容易解答.要证明这些曲面中每两个不同胚，那就需要用同调群的概念与它的拓扑不变性了.

上面所列举的曲面还不能包括所有的曲面.我们所得着的，到现在为止，只是无边缘的而且由有限个多边形所连接成的闭曲面.此外还有无穷的曲面，需要无穷多个多边形拼嵌而成.例如欧几里得平面与欧几里得空间中的单叶双曲面.

3　同痕，同伦，同调

闭曲面是流形的特款.流形的不同胚的证法，粗浅地说，是根据流形上的低维域的一种分类.现在我们用曲面上的曲线这个最简单的例子来说明.

我们先研究无重点的而且有确定的流向(Durchlaufungssinn)的闭曲线——定向圆的拓扑象.要达到一个曲面上的这种曲线的分类的目的，我们必须先规定这种的两条曲线何时等价.我们说：若是两条曲线 a 与 b 中的一条能在这曲面上绵续的变状成另一条，这两条曲线就等价.若是在把 a 绵续的变状成 b 的任一瞬间，a 所变成的曲线都无重点，这变状就是所谓同痕(Isotop)变状，而且 a 与 b 就说是同痕.例如环面上的两个同样定向的经圆同痕，图 20.24 中的曲线 Ⅰ 与 Ⅱ 同痕，Ⅰ 与 Ⅲ 却不同痕.同痕变状难以用算学的方式处置；所以与同伦(Homotop)变状比较，同痕变状将来不占重要的地位.

在 a 经过同伦变状变成 b 的历程中，a 所变成的曲线无须都无重点，其实，

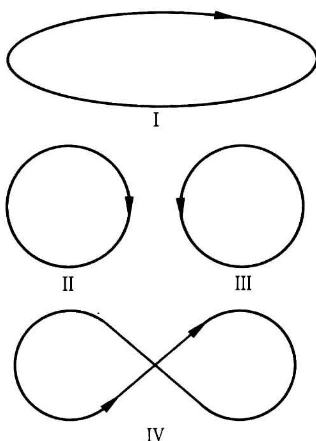


图 20.24

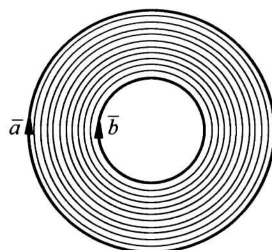


图 20.25

还可以任意自相变切. 我们也无须假设 a 与 b 都无重点. 若是 a 经过同伦变状变成 b,我们就说它们同伦,或者更正确地说,自由同伦. 同痕的曲线当然也同伦. 图 20.24 中的四条曲线的每一条都能在平面上缩成一点,都是所谓零伦(Nullhomotop) 的曲线,所以它们在平面上都同伦. 环面上的一个有确定的流向的经圆与流向相反的经圆不同伦,一个经圆与一个纬圆不同伦,所有的经纬圆都不同伦.—— 我们也能不引用变状,界说同伦的曲线如下:设 a 与 b 是曲面 $\tilde{\delta}$ 上的二曲线. 若是平环(图 20.25) 能绵续地(无须拓扑地) 换成 $\tilde{\delta}$ 上的一个点集,而它的定向的边缘圆 $\overline{a}$ 与 $\overline{b}$ 分别换成 a 与 b, a 与 b 就说是同伦地曲线. 在这种情形下, $\overline{a}$ 经历同心圆变成 $\overline{b}$,就相当于 a 同伦的变状成 b;而且,反之,在 a 同伦地变状成 b 的时候, a 经过一个"广义的"(Singulär) 平环,即平环的绵续象.

这种变状很自然地引出一种推广,即引出闭曲线的一种极粗浅的但极重要的分类法,把闭曲线分成同调类(Homologieklasse). 我们不用平环 —— 在拓扑学中,等于穿了两个洞的球面 —— 却用穿了两个洞的亏格(Geschlecht) h 的能定向的曲面(图 20.26 中表出 $h = 1$ 的专款) 替代,把它绵续地(正变换单值,逆变换无须单值) 换成 $\tilde{\delta}$ 上的一个点集. 若是这种变换存在,使洞边缘 $\overline{a}$ 与 $\overline{b}$(如图中所确定了的方向) 换成给定的二曲线 a 与 b, a 与 b 就说是互相同调(Homolog). 例如图 20.27 中所表出的曲面 $\tilde{\delta}$(安装上了三个环柄的球面) 上的曲

线 a 与 b 就同调；因为它们把 $\tilde{\delta}$ 分成两个穿了两个洞的环面，而且组成这样的每一环面的边缘．我们将来证明，同调类组成一 Abel 群，即一维的同调群．一维同调群是曲面 $\tilde{\delta}$ 的拓扑不变性．利用它，我们才能证明前节中所列举的曲面都不同胚．例如，球面只有一个同调类，投影平面有两个，而其他闭曲面的同调群都是无穷群．我们如此粗略地所指出的概念，当然还需要正确的定义；所叙述的定理，也当然需要证明．

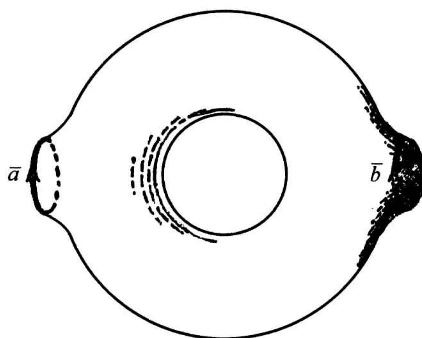


图 20.26

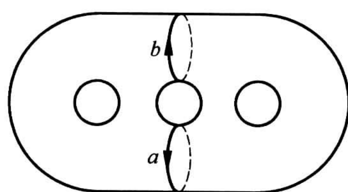


图 20.27

4　多维流形

曲面的不同胚的问题虽已能完满解答，三维的与更高维的同样的问题却还未解决．不同胚的三维空间，我们已不能完全列举，更不必说更高维的了．多维空间的不同胚问题的讨论，自然与直觉离开更远．但这问题不仅自身很有兴趣，它还在微分方程式论与两个复变的函数论中出现．虽然所有的能定向的闭曲面都是函数论中的 Riemann 曲面，三维空间中也有两个——投影空间与球式空间（sphärische raum）——在拓扑学以外的算学中，占重要的地位．

（实素的）投影空间的产生，是由于要用新点补充欧几里得空间，使投影变换在新点集中成为一一对应的变换．我们所习用的方法，是用一个假平面——在投影变换下，欧几里得空间中的没影平面（Fluchtebene）的像平面——把欧几里得空间封闭成投影空间．投影空间与欧几里得空间不同：投影空间是闭空间，能由有限个四面体拼嵌而成．

球式空间也是用新点封闭欧几里得空间而成的．不过此处加入新点的目的，是要使等角变换（konforme abbildung）——欧几里得空间中的球变换（Kugelverwandtschaft）——在球式空间中成为一一对应的变换①．球变换中之一

① F. Klein, Höhere Geometrie(Berlin 1926)50; W. Blaschke, Differentialgeometrie I(Berlin 1921)40.

是倒半径变换(abbildung durch reziproke radien),而且欧几里得空间经过这变换时,单位球(反演球 Inversionskugel)的中心无象点.我们只要用这中心的象点(不用一个假平面的所有的点)封闭欧几里得空间就得着这球式空间.这与封闭实数平面成实数球面相同.

其他多维的流体,在算学中出现的,常常不是点流体,而是另一类的元素所组成的.我们已经知道一个非点所组成的二维的流体:三维的欧几里得空间中所有通过一点的,未定向的直线所组成的集合.我们能把这流体换成投影平面,使邻近的直线换成邻近的点.一机械系统的所有位置组成另一类流体,我们举平面双摆(ebene doppelpendel)这最简单的机械系统为例.平面双摆由两条刚体杠杆 l_1 与 l_2 组成:l_1 的一个端点悬在点 A 处,它的另一端点 B 与 l_2 的一端点用一个圆关节(Gelenk)相连(图 20.28);它们在平面中自由运动,不受别种限制.这两条杠杆与垂线所成的两个角 φ 与 ψ,可以用来确定这双摆所有能占有的位置;而且,若是 m 与 n 都是整数,(φ,ψ) 与 $(\varphi+m2\pi,\psi+n2\pi)$ 确定它的同一位置.所以要表出这双摆的所有位置,我们可以用 $\varphi\psi$ 平面中边长 2π 的一个正方形,而且叠合这正方形的每两条相对边.因为这样的正方形也代表环面,我们能说这双摆的位置与环面中的点成对应,而且环面上邻近的点相当于双摆的邻近位置.双摆的一循环运动 —— 回到起点的运动 —— 所以相当于环面上的一条闭曲线.

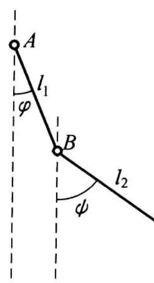


图 20.28

若是用球关节代替圆关节,我们便得着球面双摆(sphärisches doppelpendel)它的所有位置与两个球面上的点耦成对应,而且邻近的位置与邻近的点耦成对应.我们可以用四个参数,例如两个球面的经纬度,来确定这双摆的每一位置,所以球面双摆的"位置空间"是一个四维流体.

我们现在要说明三维的(实素)投影空间中所有的定向直线,也组成一个四维流形,而且它同前一个例子一样,可能看做是由两个球面的点耦组成的,所以与前一个流形同胚.设想这实素投影空间浸没在复素投影空间(它的点坐标是四个不全等于零的复数的比:$x_1:x_2:x_3:x_4$)之中.第一步我们证明定向的实直线的集合与零件

$$x_1^2+x_2^2+x_3^2+x_4^2=0$$

(它的点全是复数,那就是说,它的点的坐标是复数)的点成拓扑对应如下:每一条实直线与零球相交于两个不同的相配复点 P 与 $\bar{P}$;而且,反之,零球上的每两个相配复点确定一条连接实数 g.设 P_1,P_2,P_3 是 g 的三个实点.g 的一个定向确定这三点的一个循环次序;假设 P_1,P_2,P_3 代表这次序.因为

$$(PP_1P_2P_3) = \lambda \text{ 与 } (\overline{P}P_1P_2P_3) = \overline{\lambda}$$

是相配复数，不是实数，其中恰有一个，其虚数部分是正数；假设就是 λ. 又因为

$$(PP_2P_3P_1) = \frac{1}{1-\lambda} \text{ 与 } (PP_3P_1P_2) = \frac{\lambda - 1}{\lambda}$$

λ 的这性质不因 P_1, P_2, P_3 的循环置换(Vertauschung) 而改变，所以这定向的直线 g 唯一地确定 P 这个点. 又因为

$$(PP_1P_3P_2) = \frac{1}{\lambda}$$

的虚数部分是负的，相反定向的直线 g 确定点 $\overline{P}$. 定向实线与零球的点间因此建立了一个一一对应的关系.

现在第二步，再用"左""右" 两个二次列线(Schar) 的复直线产生这零球①. 设 l 是左二次列线中的一条直线，r 是右二次列线中的一条直线. 给定了零球的一点 P，左(右) 列线中恰有一条直线通过 P，与 $r[l]$ 交于一点 $P_r[P_l]$，所以每一点 P 恰唯一地确定两个点，一点在 r 上，一点在 l 上；反之亦然.

所以要证明投影空间中的定向直线与两个球面的点耦成一一对应，只要证明 $r[l]$ 的所有点，那就是说，一条投影直线的所有(复) 点，组成一个实数球面. 这是很显然的. 假设投影直线上有一投影坐标系 $\mu_1 : \mu_2$. 直线的点相当于 $\mu_1 : \mu_2$ 所经历的所有可能的复数(包含 ∞ 在内)，所以恰组成一个实数球面.

如同球面双摆的位置空间一样，实素投影空间中所有的定向直线的流形也与两个球面的点耦成对应，邻近的直线相当于邻近的点.

不只是力学系统的位置组成多维流形，它的运动状态(Bewegungszustand) 也组成叫做相空间的多维流形. 所谓一个质点的运动状态，就是这质点在空间中的位置，速度的量与方向. 我们举下列的例子. 设重力场中有一个质点，被限制在一个固定的球面上运动；而且它的能量 —— 动能与位能的总和 —— 在任何时间都等于一常数；而且这常数又如此之大，使它在运动历程中能走到球面的最高点. 这质点的每一运动状态可以由它的球面上的位置(这需要两个标志) 与它的速度的方向(这是第三个标志) 确定. 速度的量已可以从预先给定的能量与位置断定，所以并不必要. 所以运动状态与球面上的定向的线素(Linienelement) 成一一对应：相空间所以是一个三维空间，而且可以变成三维的投影空间. 运动状态因此相当于投影空间中的点，邻近的运动状态相当于邻近的点.

邻域，我们起初把图形看做是欧几里得空间中的点集. 后来把 Möbius 带封闭成投影平面时，我们已经丢开了这种朴质的概念，而且还另下了曲面的定义，

① F. Klein, Höhere Geometrie(Berlin 1926) § 45，或 E. Study, Geometrie der Kreise und Kugeln, Math. Ann. 86(1922)，或 L. Bieberbach, Höhere Geometrie(Leipzig 1933).





说它们是由多边形拼嵌而成的.其实,把二维流形浸没在空间中,虽有时还可以说有直觉作根据;但是那些不由点组成的更高维的流形也要用点集代表,也要把它们浸没在多维的欧几里得空间中,在一般情形下,反是既不自然又不适当的办法.所以讨论到这种流形的时候,我们更应该丢开那包含的空间.我们还应该注意,我们此前说明拓扑变换与同胚这两个概念,还是根据于图形浸没在空间中这种看法.这是因为在本书起始的时候,我们还是在空间中取一坐标系,用表示变换的坐标函数(Koordinatenfunktion)的绵续来界说变换是绵续的.但是我们在讨论双摆位置的流形的时候,我们并没有用坐标函数,却也说明了这个流形与环面相同.所以我们必须另想办法,把握着绵续的要素,不再利用空间中的坐标与坐标函数.

现在,我们试想一想,位置空间或相空间与欧几里得空间中的点集究竟具有什么共同的性质,使它们间能有连续变换存在?在位置空间或相空间中,我们直觉地知道什么元素(位置或运动状态)与某一元素邻近,什么子集——当然有各种不同的选择——组成某一元素的邻域(Umgebung).一个换邻域成邻域的变换才叫做绵续的变换.球面的点集与环面的点集的纯数(Mächtigkeit)相等,所以这两个点集成一一对应.但是它们不同胚,那就是说,它们间没有一个一一对应的变换,使邻近点换成邻近点.所以邻域的规定,是使算学的集合能成为拓扑学所研究的对象的,或者说,使算学的集合能有空间性的最低的条件.如何规定邻域,我们可举下列的空间为例.在欧几里得的空间中,一点的每一邻域,必须包含那围绕这点的一个球体中所有的点;邻域的大小,视选择的球体的大小而定.在球式空间中,一个点集必须包含有一个够大的球面外所有的点,才是这新加入的假点的邻域.在投影空间中,给定了一条直线,可以取一个细狭的单叶双曲体围绕这直线,以这直线为轴.一投影直线的邻域必须包含以它为轴的,够细狭的一个单叶双曲体之内的所有直线.

我们用邻域概念做绵续性与拓扑变换的基础,然后我们所要讨论的图形才脱离了那包含的直觉空间,因而才能与其他有空间性的对象,例如曲面等的地位.我们从此就要把空间看做是几何学中的绵续点集.邻域概念对于空间之成为空间的关系的密切,因此 也更为明显.我们将要认识的许多有意义的概念与定理,就都与距离(Entfernung)、平直线(Geradlinigkeit),甚至于空间的维数都不发生关系.若是要把一个点集当做空间,我们必须规定其中每一点的邻点.一个点集只要有了邻域的规定,无须别种限制,就叫做邻域空间(Umgebungsraum).

有了邻域概念,我们才能建立最广义的空间概念的界说,那就是说,才能把最广义的空间概念建立在集合论的基本概念之上,使这不依赖于模糊的直觉.而且如此得着的概念的普遍性最可以发挥广大的作用:只要一个算学的集合能作一种邻域的规定,使邻域适合某些公理,我们就能把这集合当做一个空间(最

广义的),而且能在这一空间中运用对于任意的邻域空间所推出的概念与定理.

我们现在要用综合几何学做模范,要采用一种邻域理论,把这章中所略述的概念与定理严密地建筑起来;因此我们必须从新起始,完全脱离直接的直觉.在开始数章中,只偶然有些例子表明普遍的讨论与几何问题间的关系.虽然如此,下一章却也不是完全讨论最广义的邻域空间.其实我们所注意的,差不多都是十分特殊的邻域空间,即我们已提起过的复合形.至于介乎我们所采用的最广义的邻域空间概念与复合形概念之间的,还有集合论的拓扑学,包括拓扑空间,它的点集与维数论,我们都不加以讨论.





希尔伯特谈拓扑①

第二十一章

在研究投影几何学时,我们已经看到了不用比较长度和角仍能描写具有明确的几何特征的现象.拓扑学所研究的几何事实,甚至不包括直线和平面的概念,只包括图形各点之间的连续连通性.设想有一个用可随意变形但不得撕破或粘连的某种材料制成的图形.我们将探讨经过这种变形而不变的属于图形的性质.举例来说,球面的拓扑性质就和椭球面、立方体、四面体的相同,另一方面,球面和环面的拓扑性质不同,因为显而易见,一球面若不许撕破也不许粘连,不会变成环面.

在几何学的发展史上,拓扑学问题的出现比投影几何学还要迟,这是自然的.拓扑学的问题直到 18 世纪才开始研究.晚些发现,拓扑学的一些定理,纵然在外表上有不确定性,却同许多最精确的定量的数学结果发生联系.这里是指代数学、复变数函数论和群论.时至今日,在数学各分支中,拓扑学的研究是最为丰富多彩、最有收获的一种.

今后我们限于研究关于三维空间中的曲面的几个问题②.首先从拓扑学上最容易探讨的曲面 —— 多面体入手.

---

① 本节选自:希尔伯特 D,康福森 D.直观几何下册[M].王联芳,译.北京:高等教育出版社,1964.

② 要想进一步通晓拓扑学的基本概念,可参考 П. Александров, Einfachste Grundbegriffe der Topologie.

1 多面体

所谓多面体,是说任意多个多边(角)形的组合,这种组合适合两个条件:(1) 在任一棱相遇的多边形恰恰有两个(作成二面角);(2) 从任一多边形出发,跨过一些棱,能到任一别的多边形上去.

最简单而又最重要的多面体要算简单多面体.所谓简单多面体,是说经过连续变形能变作球面的多面体.正多面体就是简单多面体的例子.不久我们会看到,也有好多多面体不是简单多面体;也就是说,有不能变作球面的多面体.

我们知道,正多面体有一特性:没有凹进的棱.由此引申得到,正多面体是凸的."凸"的意思是说整个多面体居于每一面的同一侧,因而我们可以拿任一面作底把它放在桌面上.凸性不是拓扑性质,因为多面体经过按照拓扑观点是非本质的交换能变为非凸的多面体.虽然如此,多面体的凸性却蕴涵着一个拓扑性质.因为,稍加思索即可明白:任一凸多面体一定是简单多面体①.

简单多面体的顶点数、棱数和面数之间有一个很重要的关系.这个关系是欧拉发现的,因而叫做多面体的欧拉定理(公式).设 E 是多面体的顶点数,K 是棱数,F 是面数,欧拉定理断言,在所有的简单多面体中

$$E - K + F = 2$$

姑以下面举出的几个多面体为例来验证这个奇妙的定理.

四面体: $E - K + F = 4 - 6 + 4 = 2$

立方体: $8 - 12 + 6 = 2$

八面体: $6 - 12 + 8 = 2$

为了证明欧拉定理,我们在一平面上作一简单多面体的象,即所谓多面体的平面网,其法如下:从多面体中任意挖掉一个面,把其余的各面一齐挤压到这个面上来.在这样变形时,还能办到使各个面依然是以直线段为边缘并且顶点的总数也不变的多边形(但是当然不能作到使每个平面多边形都与原多边形合同).这个多边形组称为多面体的平面网.

平面网所含的顶点数和棱数跟多面体的相同,只是面数少了一个.现在让我们对平面网作一系列的变换,使网的构造简化,但不变更 $E - K + F$ 这个数.

首先,凡遇到在平面网中边数大于 3 的多边形,就联一对角线.这样将多出

---

① 凸和非凸的多面体有一个突出的差别:一方面任一凸闭多面体都是刚性的,另一方面,有的非凸闭多面体,它的面可以作相对运动.但一直到现在还没有人能用极限的方法从多面体的刚性直接推出凸闭曲面的刚性.—— 原著者注

借助从多面体出发的极限过程以研究曲面的度量的方法,近几年来已为 А · Д · 亚历山大洛夫所发展(参考他著的《凸曲面的内蕴几何学》,Гостехизат,1948).А.В.Погорелов 根据亚氏方法,不用解析方法而用几何方法证明了凸闭曲面的刚性(参考他的"Однозначная предельность Выпуклых ловерхностей",Труды Математического института пм.В.А. Стеклова АН СССР, Выпуск XX IX,1949).

亚氏书有吴祖基的译本,科学出版社出版.



一个面和一个棱,但顶点数未变,因此 $E - K + F$ 不变(图 21.1).继续用这个办法,一直到各面都成为三角形为止.

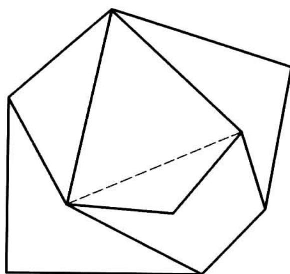


图 21.1

如果沿着作好的三角形网的某一个棱添上一个三角形,使新三角形有一棱是原三角形网中的一棱(图 21.2).这样顶点数和面数各多了一个,棱数多了两个,因而 $E - K + F$ 仍不变.此外,若在边缘陷入的部分过二顶点联一线段(图 21.3),这样虽然多出一个三角形,因而多出一个面,但棱数多一而顶点未变,因而 $E - K + F$ 仍不变.

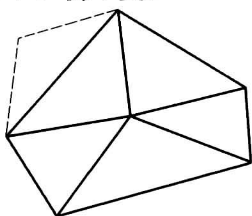


图 21.2

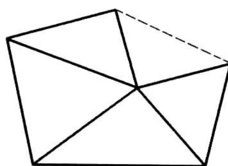


图 21.3

容易看出,任一三角形网都可以经过多次重复上述两种手续从一个单独的三角形产生出来,所以 $E - K + F$ 之值对于任一三角形网,从而对于任一平面网,和一个三角形的相同:$E - K + F = 3 - 3 + 1 = 1$.不过,平面网的顶点数和棱数和简单多面体的相同,仅面数少一;因此对简单多面体必有

$$E - K + F = 2$$①

欧拉定理提供了正多面体只有 5 种类型这一事实的一个新颖而简单的证

① 庞加莱将欧拉定理推广到 n 维空间.代替顶点棱和面的是 $0,1,2,\cdots,n-1$ 维的形象.设这些形象的个数分别用 $N_0,N_1,N_2,\cdots,N_{n-1}$ 表示,则对于和简单多面体相当的流形有下列关系式

$$N_0 - N_1 + N_2 - \cdots = 1 - (-1)^n$$

当 $n = 3$ 时,这就归结为欧拉定理.

明在我们所考虑的正多面体中,遇于一顶点的面数设为 n,因而在这点相遇的棱数也是 n.又假定 E,K,F 的意义同上,那么通过所有顶点的棱数等于 nE.但每一棱已计算了两次,因为每一棱连接二顶点.所以

$$nE = 2K$$

假定多面体的每一面由 r 个棱所围成,则 rF 是包围多面体各面所有的棱数.这里每一棱又计算两次,因为每一棱介于二面之间,所以

$$rF = 2K$$

将上二式代入欧拉公式,得

$$\frac{2K}{n} - K + \frac{2K}{r} = 2$$

此式又可以写作

$$\frac{1}{n} + \frac{1}{r} = \frac{1}{2} + \frac{1}{K}$$

按 n 和 r 的含义,每个数至少要等于3.但它们不能同时大于3.假定 n 和 r 都大于3,则应有

$$\frac{1}{K} = \frac{1}{n} + \frac{1}{r} - \frac{1}{2} \leqslant \frac{1}{4} + \frac{1}{4} - \frac{1}{2} = 0$$

这是不可能的.今设 $n = 3$,则得

$$\frac{1}{K} = \frac{1}{r} - \frac{1}{6}$$

由此可见,对于 $n = 3$,r 只能取3,4,5三个值,从此得出 K 的值等于6,12,30.因为方程对于 n 和 r 对称,故当 $r = 3$ 时立即得出 n 的对应值.这样说来,我们一共得出六种可能的情形,其中有两种是相同的,即 $n = 3, r = 3$.所以只剩下5种不同类型的多面体,并且它们确实在正多面体中实现①.

现在转向非简单多面体.今以棱柱块为例(图21.4).这个棱柱块乃是一个四棱柱,中间挖去一个洞,洞的形状仍是一个四棱柱,这两个四棱柱的边分别平行.又,二棱柱所共有的底面向外倾斜,如图所示.这个棱柱块是不能变形成球面但可以变形成环面的多面体②.若在棱柱块上仿照上述方法挖几个洞,可得出其他类型的多面体(图21.5).

① 相仿地,欧拉定理的庞加莱推广提供了决定高维空间的正则胞腔的一种方法.
② 棱柱块也是拓扑正则的.



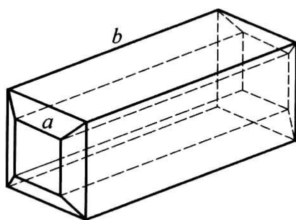


图 21.4

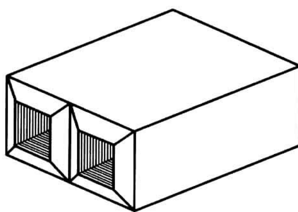


图 21.5

要想统观种类繁多的多面体的全貌,我们给每一多面体一个确定的数 h,所谓多面体的连通度.今考虑由一个多面体的棱组成的闭合而不自交的折线(闭棱折线).如果每条这样的闭棱折线都把多面体的表面部分为分离的两片,则命此多面体的连通度 $h=1$.显然,所有简单多面体的连通度都等于 1,因为球面被自身上的任一闭曲线剖分为两片.反之,也容易看出,连通度等于 1 的多面体总可加以连续变形,使之成为球面.由于这个缘故,简单多面体也叫做单连通的多面体.

另一方面,也有在棱柱块上的一条闭棱折线(例如图 21.4 中的正方形 a)不能把棱柱块剖分为两片的情形.凡具有这种性质的多面体我们给它大于 1 的连通度.为了确定它的连通度,应同时考虑连接第一次所作的棱折线上两点的所有的棱折线(不一定是闭合的).

如果任何这样的一对棱折线剖分多面体的表面为二,我们就说它的连通度 $h=2$.若剖分不开,再重复上述步骤.一般地,我们有下面的定义:

我们说一多面体的连通度为 h,如果在该多面体上可以找到有一定的顺次的 $h-1$ 条,但不是 h 条棱折线,剖分不开多面体的表面①.此处约定第一条棱折线必须是闭合的,而每一后继的棱折线则连接前一条棱折线上的两点.

在棱柱块上,从图 21.4 上容易看出,有两条这样的棱折线(指正方形 a 和梯形 b),不能把多面体的表面剖分为两片.所以该多面体的连通度至少是 3.下面我们要证明,它的连通度恰好是 3.

现在发生这样的问题:能不能把已证明对于单连通的多面体成立的欧拉定理推广到有任意连通度 h 的多面体上去呢?我们不敢奢望定理一字不改,因为那个定理的证明原来借助于“平面网”,而平面网的结构,显然只有对单连通的多面体才有可能.不过可以证明,一般地应用

$$E-K+F=3-h$$

当 $h=1$ 时,这个方程就是前面证过的欧拉公式.现以上述棱柱块为例,显然它有 16 个顶点,32 个棱,16 个面,因此有

① 这是说,多面体上任意两点总可能连以在多面体上的,但不同任一棱折线相交的曲线.

$$16 - 32 + 16 = 3 - 3 = 0$$

从而可知,棱柱块的连通度恰好等于 3.这个一般的欧拉定理对于决定多面体的连通度甚为方便.应用这个办法只需数一数顶点、棱和面的个数,而不必把棱折线一条条地检验.

2 曲面

在上一节我们看见过,简单多面体可以经过变形而成球面,棱柱块可以经过变形而成环面.相仿地,一些更复杂的拓扑图形也可以用像多面体那样的图形来代替.用这种方法,拓扑图形的理论归结到由简单元素组成,而组成的方式易于描述的图形的研究.这种所谓"组合拓扑学"的研究方法有很大的优点:它可以立刻推广到三维以上的空间去;因为每一多面体的结构,用系统的组合规则而不假借直观力之助,即可安全描写出来.

话虽如此,可是直接拿曲面作基础,更容易为直观接受.例如,球面的结构比多面体的简单,环面的结构比棱柱块的简单.因此我们想把多面体连通度的概念推广到任意曲面上去.

对于球面,连通度 h 应该定为 1,对于环面,则 h 是 3.要得到更高的连通度,可用容易揉捏的物质做成球,捏扁,再开上几个洞(图 21.6).这种曲面我们称之为"圈饼"①.可以证明,有 p 个洞的圈饼,其连通度 $h = 2p + 1$.图上画出了几个连通度不同的圈饼.剖分不开曲面的曲线的数目是0,2,4,6,圈饼的连通度则分别是 1,3,5,7.容易看出,若是再任意作一条连接截口上两点的曲线,就能把曲面分成两片了.

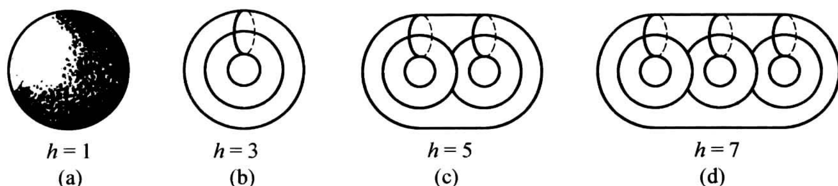


图 21.6

在一般曲面上选择曲线比在多面体上自由,因为在多面体上只能由棱折线里挑选.因此我们可给曲面的连通度下不同的定义,下面的是其中的一个:

在连通度为 h 的曲面上,可作 $h - 1$ 条闭曲线,剖分不开曲面;但任一组 h 条的闭曲线至少把曲面分成两片.

图 21.7 上表示对于 $h = 1,3,5,7$ 各种情形的曲线组.

---

① 圈饼是德国一种绞环形的脆饼干.—— 原译者注

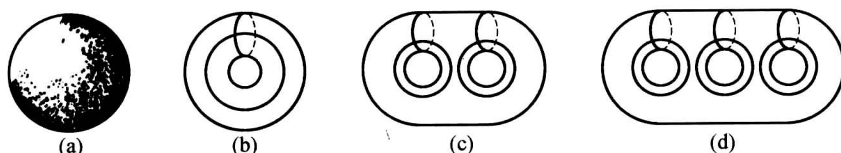


图 21.7

我们还可以加入另一个条件,就是要曲线通过曲面上任意选定的某一点.这样就得到曲面的“典型截口组”,这对于某些目的甚为方便.图 21.6,21.7,21.8 是典型截口组的例子.

另一方面,若限制闭曲线组互不相交,上述结果须加修改.对于奇数连通度的曲面可以证明:

在连通度 $h = 2p + 1$ 的闭曲面上,有一组剖分不开曲面的 p 条,但不多于 p 条互不相交的闭曲线.

这个定理的真实性可从图 21.7 看出来.

直到现在,我们仅仅研究了有限闭曲面(没有边缘).其实连通度的概念还可以推广到更一般的情形上去.首先,让我们假定曲面是有限的,但有若干个闭边缘曲线.假定这些曲线不跟自己相交,也不彼此相交.图 21.8 上画出一些这类的曲面片.还有一些这类的曲面可从图 21.6 和 21.7 的一些闭曲面得到,办法是把它们想象为中空的,并且挖开几个洞(参考图 21.9).①

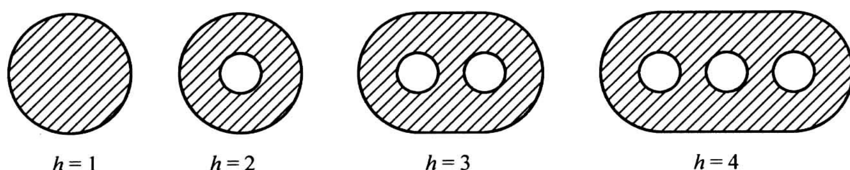


图 21.8

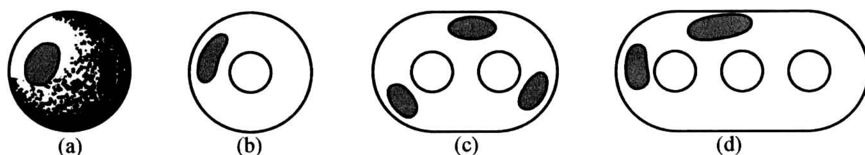


图 21.9

在有限有边缘的曲面中,同以前一样,也用一组曲线来定义连通度,只有一点不同:第一条曲线不是闭合的,而是联结边缘上两点的曲线.至于后继的曲线,可以以这组中前面的那些曲线上的两点为起讫点,也可以联结边缘上的两

① 同图 21.8 的曲面相反,图 21.9(b),(c),(d) 不能用一张纸剪几个洞经过变形而成.这种区别在几何函数论中甚为重要(简单域和非简单域).

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

点.按照这个定义,即图 21.9 上各曲面的连通度应该依次为 2,3,7,8.

用闭曲线来规定连通度的办法,不能直接推广到有边缘的曲面上去.

现在我们来考虑伸展到无穷远的曲面,而且它可以有边缘,也可以没有边缘.这样曲面的拓扑结构要看曲面是安装在度量空间里或安装在投影空间里而定.

在度量空间里我们姑限于讨论曲面的有限点.这样,就可设想空间好像被包围在一个非常庞大的球面内,因而我们的曲面可换成它在球面内的那部分.这正是属于有限又有边缘的曲面的问题,因而便可应用前面已讲过的理论.①

在投影空间里,情形完全不同.我们把每一直线当做是闭曲线,其上有一无穷远点,曲线的二支就在这里汇合.除此之外,这一无穷远点认为同时属于前一条直线的所有平行线.按照这些规定,则作为一个整体的投影空间也是在它自己的诸无穷远点处连通的.曲面包含一无穷远点的条件是:曲面上有一路线,此路线越来越接近于这个无穷远点所属的某一直线.这里不一定要求曲面与此相反的方向上接近于某一平行线.如果曲面同时在两个相反的方向上接近二平行线,我们应认为曲面在无穷远点连通.但若曲面仅沿某一方向接近于一直线,而不在相反的方向上接近于它的平行线,则直线上的无穷远点是一边缘点.其次,假如曲面有一条伸展到无穷远的边缘曲线,则这条曲线必在无穷远闭合;也就是说,曲线若不是在相同的方向上或相反的方向上渐近地接近于二平行直线,就是包含无穷远线的一部分.因为一非闭曲线不能作成曲面的边缘.例如在投影空间里,像图 21.10 上标出的介于一直线和一折线之间的部分平面不跟平面的其他部分分开,因为比如从 A 到 A' 可以取道无穷远.但在度量空间里,这一部分平面就好像以一条闭曲线为边缘似的.

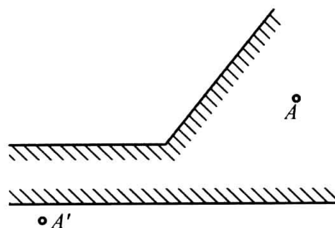


图 21.10

相当的说法也适用于作为一个整体的平面.度量平面以无穷远直线为闭边缘曲线,因此它同圆面拓扑等价.另一方面,投影平面是闭曲面.我们也可给投影平面作一个简单的拓扑模型.在那里通过叠合半球面边缘上大圆的每对对径

① 必须假定球能选择得这样大,使得任何更进一步的扩大都保持曲面在球内的部分的拓扑结构不变.不难找到不适合这个条件的例子.我们绕一个平面正方形点格的各点作无公共点的小圆,再从平面上去掉各圆的内部,则得一平曲面,对于这个曲面在任意已给球内的部分,容易计算出它的连通度.但是,显然这个连通度,在保持球心固定的条件下,随着球的增大而增大.

点而建立投影平面的一对一映射. 当然也可用圆面代替半球面,因为圆面总可以连续变形而成半球面. 现在让我们把圆面变为正方形面. 于是投影平面同正方形面拓扑等价(图 21.11),但此正方形边缘上的每对对径点(如图 21.11 上的 A 和 A' 等)须视为叠合起来的. 在这个模型中,同投影平面闭曲线对应的不仅是正方形上的闭曲线,还有一切连接一对对径点的曲线(如图 21.11 的直线段 AA' 等等).

此刻我们要考虑由图 21.11 的办法直接引出来的其他类似的做法. 首先我们还是从正方形面或矩形面开始,但须按照图 21.12 表示的那样迭合对径点. 于是我们又得到一个闭曲面的模型. 这一次我们容易从模型恢复所代表的曲面. 首先把矩形变曲成圆筒形(图 21.13,21.14),并将 1,2 两边这样粘牢,使得二边上每对叠合点重合. 这时 3,4 两边变成了圆周,再把这二圆周按同一规则粘在一起(图 21.15). 最后得到环面,这时原来的矩形面的边缘变为环面的"典型截口组",其中每一截口相当于矩形面的一边对边(看图 21.16 和 21.17(b)). 反过来,从环面开始,将环面沿典型截口组切开,则得一个图形,它与对边按上述规则叠合的矩形面拓扑等价. 这一作图方法也适用于所有的"圈饼". 对于连通度为 $2p+1$ 的圈饼,典型截口组由 $2p$ 条曲线组成;沿这些曲线剖开,得到一个有 $4p$ 条边的且各边按一定规则成对地叠合的多边形. 图 21.17 和图 21.18 演示 $h=5$ 和 $h=7$(即 $p=2,p=3$) 的情形的分步做法.

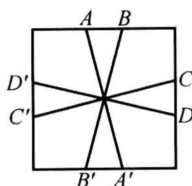


图 21.11

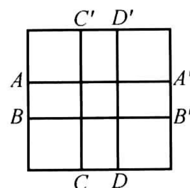


图 21.12

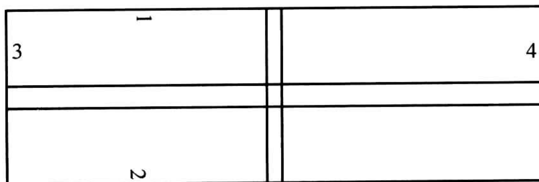


图 21.13

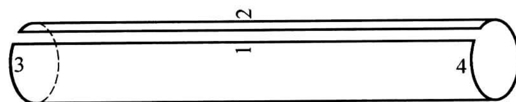


图 21.14

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

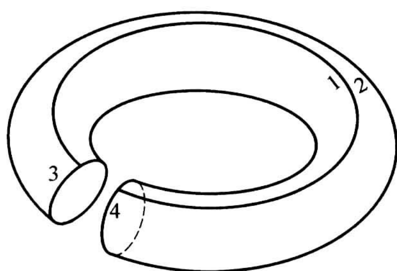


图 21.15

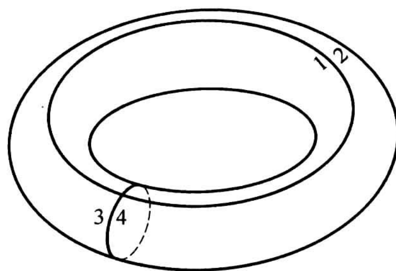
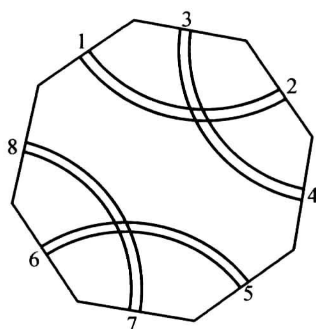
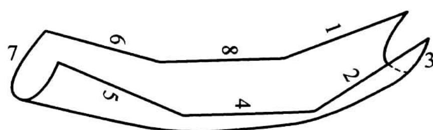


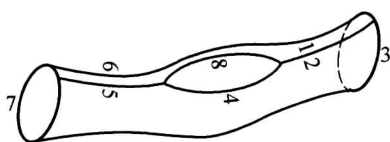
图 21.16



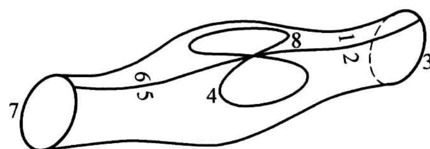
(a)



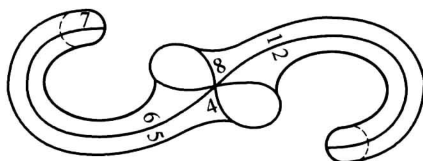
(b)



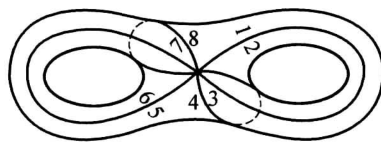
(c)



(d)



(e)



(f)

图 21.17



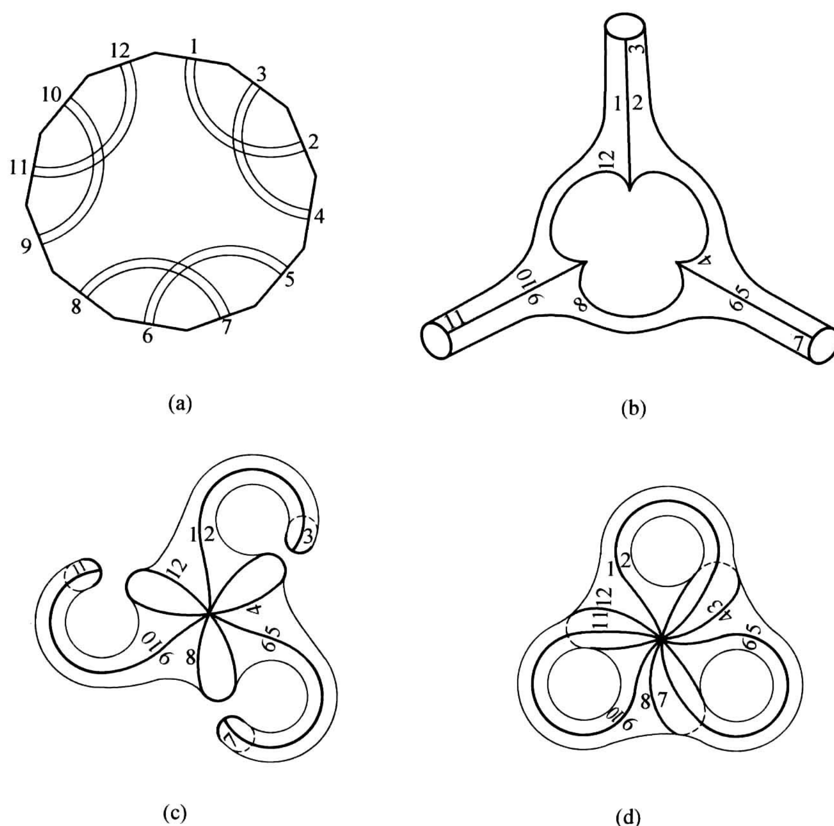


图 21.18

将圈饼映成 $4p$ 边多边形的映射，在连续映射理论和函数论中有重要的应用. 两种应用都依据这样的事实：每 - $4p$ 边正多边形都是正则在覆盖双曲平面(若 $p=1$ 则为欧氏平面).

如果把 $4p$ 边多边形的各边以不同的方式叠合，则可得到圈饼之外的许多曲面. 在下一节里我们准备讲几个这样的曲面.

3 单侧曲面

所有以前我们讨论过的多面体和闭曲面的连通度都是奇数. 于是自然会发生这样的问题：有没有连通度为偶数的闭曲面呢？换句话说，有没有按拓扑性质来说，介于球面和环面之间或者介于两个"圈饼"之间的曲面呢？

答案是肯定的. 的确，我们可作出一个多面体 —— 七面体，其连通度按多面体的欧拉定理等于 2. 这些目的，我们从有 8 个三角形面的正八面体开始(图

21.19),增添由对角线决定的三个正方形面(例如图 21.19 的 $ABCD$).用这种方法得出含十一个面的图形,它不符合从前讲的多面体的定义,因为交于每个棱的面数是三个而不是两个.现在我们去掉四个三角形:从图形的前部(按照图 21.19 的位置)去掉左上角的三角形和右下角的三角形,从图形的后部去掉左下角的三角形和右上角的三角形.这样一来,只剩下了有阴影线的四个三角形.总共我们得出由四个三角形和三个正方形组成的图形.这图形的棱和顶点是八面体的棱和顶点,但八面体的对角线却不是我们的图形的棱而是面的自交线.显然在每一棱上恰有二面相遇,而且我们可从任一面开始走起,越过若干个棱而达到任一别的面.所以这图形是多面体.由于它有七个面,所以叫做七面体.跟八面体一样,七面体也有十二个棱和六个顶点.由推广的欧拉多面体公式得方程

$$E - K + F = 6 - 12 + 7 = 3 - h$$

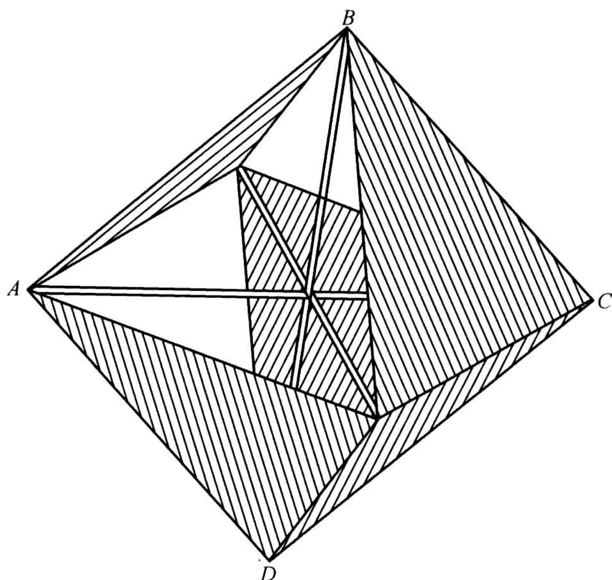


图 21.19

由此得出七面体的连通度 $h = 2$.正像简单多面体经过连续变形最后能变成球面一样,七面体经过连续变形也能变成一个简单而闭合的曲面.这种曲面称为"罗马式曲面",是斯坦纳(Steiner)研究过的.像七面体一样,这种曲面自交于三条相互垂直的线段.它的方程以直角坐标表示时是

$$y^2z^2 + z^2x^2 + x^2y^2 + xyz = 0$$

可见它是四阶的曲面.

七面体除了有偶数连通度和自交线外,还有此前我们未讨论过的另一重要性质.设想七面体是用薄膜作的.有一个会爬的小动物,例如甲虫,从曲面上的

一点 P 开始爬行.在薄膜的另一侧正对着点 P 有另一点 P'—— 如果薄膜换成原来的几何曲面,那么 P' 就和 P 重合了.可能有人会这样想,甲虫若不在薄膜的某处咬破一个洞就不能从点 P 爬到点 P'.对于我们从前讲过的球面和一切"圈饼"来说,情形确实是这样的,可是对于七面体来说就不对了.在平行于图 21.20 的平面的正方形面朝着观察者的一侧上选择一点 P,作为甲虫的出发点(图 21.20).考虑七面体上从点 P 出发越过 1,2,3,4 诸棱又回到原来正方形面的一条道路.显然甲虫从正方形面向前的一侧开始,顺着这条道路爬行,越过标号 4 的棱到达原正方形面的背后.甲虫当然要在三个地方穿过薄膜,但每次被钻穿的面乃是在七面体自交线上阻碍交通的那个面,而不是甲虫在其上爬行的面.

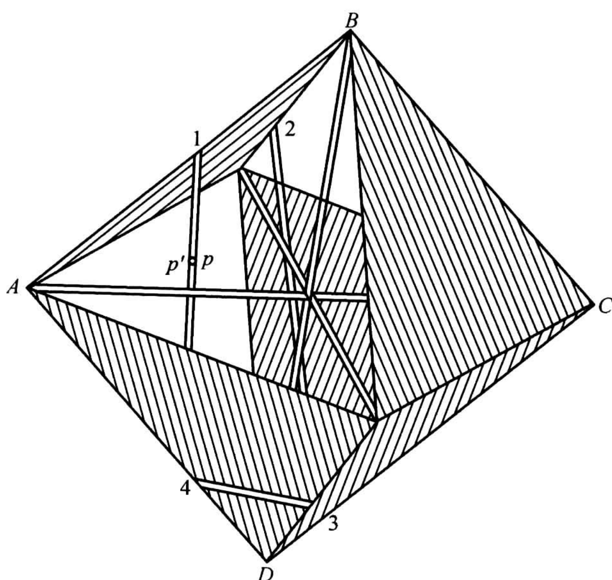


图 21.20

由于这种理由,七面体称为单侧曲面,而球面和前面讲过的各种"圈饼"称为双侧曲面.此种分类法也适用于有边缘曲面.设想有边缘曲面也是由薄膜制成的,试决定在薄膜上有没有从一侧出发到另一侧的通路,中途不许越过边缘,也不许穿过薄膜.如有这种通路存在,则曲面称为单侧的,否则称为双侧的.前面我们讨论过的所有有边缘曲面,例如圆面,都是双侧的.但是也可以举出一个比七面体更简单的有边缘的单侧曲面的例子,这就是麦比乌斯带.麦氏带可用一张细长的长方形纸条作成,如图 21.21 所示.把两边 AB 和 CD 粘起来,使 A 和 C 重合,B 和 D 重合,则得到一圆筒形带,如前所述.这是双侧有边缘曲面,另一方面,若在两端黏合之前,把一端相对于另一端扭转 180°,然后再使 A,D 重合,B,D 重合,便得到麦比乌斯带的模型(图 21.22).容易看出,这种曲面是单侧

的.因为,在两端黏合之前作一条直线,比方说是 PP',平行于带子的长边,则在两端黏合之后,直线 PP' 就变为从带子一侧到另一侧的通路 QQ'①.

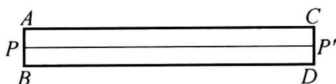


图 21.21

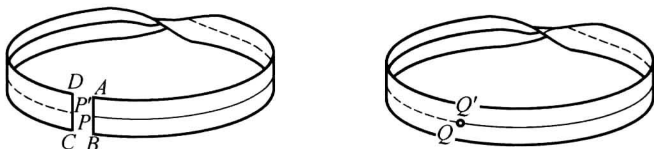


图 21.22

单侧曲面也可以用另一种重要的拓扑概念来刻画,而不必假借曲面是由薄膜制成的.设想任一给定曲面上的每一点(边缘上的点除外,如果有的话)被完全在这曲面上的一条小闭曲线所包围.现在对所有的这些曲线如此规定指向,使充分靠近的二曲线的指向相同.如果规定这样一种行进的指向是可能的,则称这样的一个规定为曲面的一个定向,而曲面称为可定向的.现在我们证明,单侧曲面不可定向.为了证明,我们考虑和曲面的单侧性等价的一条闭路.比如说,在麦比乌斯带上选好一条道路 QQ',其中的 Q 和 Q' 两点还是当做叠合的.让我们给 Q 指定一个指向.指向既经确定,在整个道路 QQ' 上就一直沿用它.这样,当动点又来到 $Q' = Q$ 时,则伴随 Q' 的指向必同 Q 的相反.假如麦比乌斯带是可定向的,这种现象便不会发生.所有别种形式的单侧曲面也和麦氏带有类似的情况.反转来也可以证明,所有双侧曲面都是可定向的.由此可知,曲面的双侧与单侧之分,和曲面的可定向与不可定向是一回事.

容易懂得,曲面之不可定向,等价于曲面上存在一条闭曲线 s,当圆心在这条曲线上连续运动时,有指向的小圆带着相反的指向回到出发点(例如图 21.22 中的曲线 QQ').在这样的曲面上,沿着曲线 s 的一侧行进,虽然没有越过曲线,也将达到它的另一侧.由于这种原因,s 称为单岸曲线,虽则一个可定向的曲面上所有的曲线都是双岸的,但单岸闭曲线的存在却是不可定向曲面的特征.曲面的单侧性同曲面上存在单岸曲线是等价的.前一性质涉及曲面在空间中的位置,后一性质涉及曲线在曲面上的位置.

七面体有自交线,这一点跟麦比乌斯带不同,那么试问任何单侧闭曲面是不是都自交呢?假如曲面只有一侧,就不可能把空间的一部分同其余的部分分

---

① 下面的两个性质也显出带与圆筒形带不同之点.第一,麦氏带的边缘是一条闭曲线,圆筒形带的边缘是两条闭曲线;第二,带沿曲线 QQ' 剪开后不分离,圆筒形带则不是这样.





开,即不能把空间分为"内""外"两部.不自交的闭曲面是不可想象的.事实上,所有单侧闭曲面都要自交.不过这句话的证明,必须另谋途径.

并不是说无论什么曲面的自交都是拓扑性质.例如,考虑图 21.23 所画的曲线绕虚线为轴旋转而得的曲面.这一曲面有自交线,即点 A 所作的圆.但曲面经连续变形,可变成如图 21.24 中的曲线所产生的回转曲面.当然这个曲面没有自交线,因此同球面等价.反过来球面也可以经过变形而变为第一个回转曲面.由此可见,自交线的存在不一定就表示拓扑性质.在这个例子里,自交线作成一闭曲线.另一方面,七面体的自交线有六个端点,即七面体的六个顶点.这些点其实应该当做奇点看待.因为曲面上任何正则点的邻域总可经过变形而成为圆面,但对七面体顶点的邻域则不可能这样作(参考图 21.19).所以七面体有六个奇点.从这里又引出一个问题:有没有无奇点的单侧闭曲面呢?

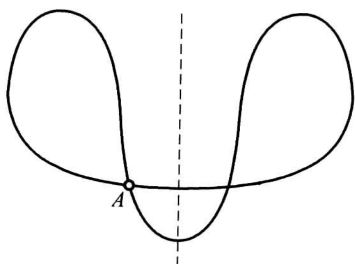


图 21.23

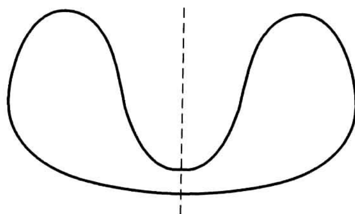


图 21.24

有的,这样的曲面第一次被克莱因发现.我们从一个两头开口的管子出发(图 21.25).前面讲过从这种管子可得到环面:把管子弯曲,一直弯曲到两端相遇,然后将两圆口粘在一起.

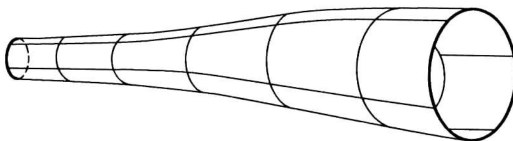


图 21.25

现在我们改用另一种方法将两端黏合.先将管子的一端捏得比另一端稍细一些,将细的一端掉过头来插入粗的一端的管壁内,使二瓶口作同心圆状(图 21.26).随后将小口放大,大口缩小,使两口重合在一起.最后把它们粘起来.这样得到的曲面没有任何奇点;它叫做克莱因曲面(也叫做"克莱因瓶"),其形状如图 21.27.这种曲面显然是单侧的,并且在细端插入管壁处沿一闭曲线自交.

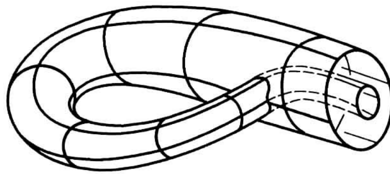


图 21.26

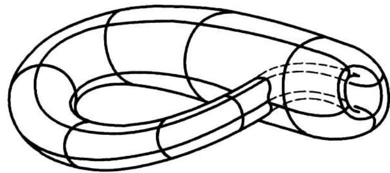


图 21.27

第一个单侧闭曲面的例子 —— 七面体,跟前面讨论过的各种双侧闭曲面不同的地方,在于七面体有偶数的连通度.那么我们可不可以抱着同样的希望,说克莱因瓶的连通度也是偶数的呢?不可以.事实上克莱因瓶的连通度跟环面一样,是 3.为了证明,我们也可以像对待环面那样,来选择克莱因瓶的典型截口组.可以选择瓶子两端的接缝作为组中的第一条闭曲线.沿着这条曲线剪开,把曲面拉直使成为圆筒形,然后选择圆筒上的一条母线作为组中的第二条曲线.沿着这条母线将曲面剪开,于是克莱因瓶就变为矩形,像环面一样.再过矩形边缘上任意两点作曲线,曲面就被分成两片了.应用连通度的一般定义于克莱因瓶,可知 $h - 1 = 2$,故 $h = 3$,这正是我们要证明的.

现在,我们已能以不同的方式叠合矩形(或正方形)的对边,得出五种不同的曲面①.下面的表 21.1 列出这五种曲面及其相应的特征.其中有关投影平面的部分,以后将予以证实.

表 21.1

	圆筒面	二条边缘曲线	$h=2$	双 侧 的
	麦比乌斯	一条边缘曲线	$h=2$	单 侧 的
	环 面	闭 曲 面	$h=3$	双 侧 的
	克来因瓶	闭 曲 面	$h=3$	单 侧 的
	投影平面	闭 曲 面	$h=2$	单 侧 的

这张表指出,麦比乌斯带的模型可以从克莱因曲面消去一组对边的叠合关

① 在投影空间里单叶双曲面看做是在无穷远连通的闭曲面.从这种观点出发,再参考这张表,读者可以决定单叶双曲面是与克莱因瓶还是与环面等价.

系而得出.因此一定能够把克莱因瓶沿着一条适当选择的闭曲线剪开,使成为麦比乌斯带.至于如何剪法,留给读者作为练习(用模型来试).图 21.28 表示用另一种方法把克莱因瓶剪成二条麦氏带.建议读者作出正方形模型的对应演变过程.

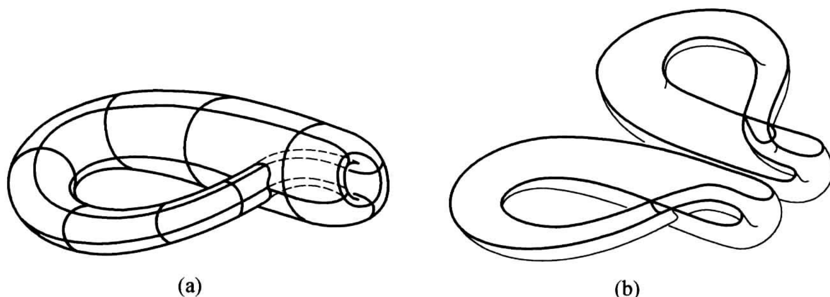


图 21.28

在我们讨论过的一些单侧闭曲面里有连通度为偶数的(例如七面体),也有连通度为奇数的(例如克莱因瓶);另一方面,我们讲过的一切双侧闭曲面的连通度都是奇数.进一步可以证明,并无连通度为偶数的双侧曲面.

正像正方形一样,一切其他的正 $4p$ 边多边形也可以用不同的叠合方式由它得出许许多多的曲面模型,其中包括有边缘和无边缘的,单侧和双侧的.如果在图 21.29 上的 AB 和 CD 是 $4p$ 边形中将要叠合的二边,则有两种可能的叠合方式:(1) 联结对应点的二直线不相交;(2)这样的二直线相交.例如在图21.29上,第一种情形由叠合 A, C 而得,第二种情形由叠合 B, D 而得.现在我们断言:如果 $4p$ 边形的任意二边是按照第二种方式对应的,则无论其他的边如何对应,所得的曲面永远是单侧的.

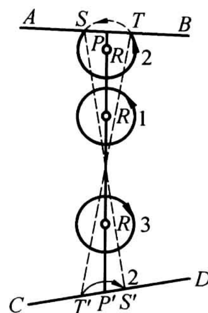


图 21.29

为了证明这句话,即证明多边形所代表的曲面不可定向.设 P 和 P' 是在叠合中 AB 和 CD 上的二对应点(图21.29),那么直线 PP' 代表曲面上的一条闭路.用 R 代表在曲面上沿这条路线运动的一点,R 先经过 P,又由 P' 回到原处.让我们给曲面上点 R 指定一个绕行指向,这个指向在 R 行进时没有任何不连续的变化.为此我们以 R 为圆心作一小圆,并以箭头表示它的指向,把小圆随着 R 连续移动.只要整个圆在 $4p$ 边形内,它就是多边形代表的曲面上一闭曲线的象.当 R 接近 P 时,只有圆弧 ST 仍留在 $4p$ 边形内.要想得到曲面上闭曲线的象,必须利用在直线 CD 上和 S, T 对应的点 S', T'.但因为 AB 和 CD 的叠合是第二类的,S 和 S' 在直线 PP' 的不同侧,T 和 T' 也如此.这样一来,在曲面上带有

绕行指向的闭曲线由二有向弧$\widehat{ST}$和$\widehat{T'S'}$代表.当 R 达到P然后从P'回到原来位置时这条曲线没有不连续的变化.随着动点 R 与 P' 的距离逐渐增加,弧 ST 逐渐消失,而弧 $S'T'$ 变为整个圆.但是这圆的指向同原来圆的相反,这就证明了曲面是不可定向的.

作为这个定理的特例,我们得到投影平面是单侧的结论.因为,在投影平面模型里,所有的对应关系都是第二类的.

反转来容易证明,如果所有的对应关系是第一类的,则模型永远代表双侧曲面.

我们已经从球的表面得出投影平面模型,另一方面,我们见过克莱因瓶同环面也有关系,虽然这种关系不同于球面和投影平面的关系.我们现在指出,事实上也可以在克莱因瓶和环面之间建立像球面与投影平面之间一样的对应关系.一般地,对任何单侧曲面,必有一双侧曲面同它对应.

要想从球面得到投影平面,必须把球面的对径点叠合起来.现在我们把类似的做法应用于环面.从环面任一母圆和圆心作环面的轴的垂线.这二直线的交点 M 叫做环面的中心(图 21.30).设 P 是环面上的任一点,则 P 对于点M 的对称点 P' 也必在环面上.环面上每对对于点 M 对称的点叫做对径点.把每对对径点看做一点,球面就变成一新的曲面 F.今证明 F 便是克莱因瓶.

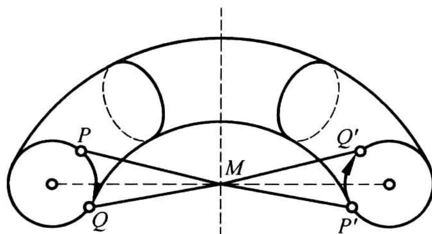


图 21.30

试考虑环面上的一母圆.这母圆对应第二个母圆,如图 21.30 所示.二母圆把环面分成两半.曲面 F 是由去掉环面的一半,并把剩下的一半的边缘圆周上的点如图上指出的方式叠合而成的.相仿地,我们曾经用半球面代替球面来构成投影平面.考虑到二叠合圆的绕行指向,可知半环面由叠合而变成克莱因瓶.

显然第二个半环面可以覆盖在第一个半环面上,使原来的每对对径点现在都重合.但是,这里必须把第二个半环面的里面翻到外面来,像翻手套一样.如果再把这两个半球面粘起来,最后就得双重覆盖的克莱因瓶①.由于这种理由,环面可叫做克莱因瓶的双叶覆盖曲面.同理,球面叫做投影平面的双叶覆盖曲

① 这种做法不能从环面的简单变形得到,必须把环面剪成两半,把一半的里面翻到外面来.

面.一般地可以证明,每一单侧曲面有一双侧曲面作为它的双叶覆盖曲面.

4 作为闭曲面的投影平面

为了决定投影平面的连通度,我们应用多面体的欧拉定理于正方形模型.通过正方形的中心点 M(图 21.31)作两边的平行线 PQ 和 RS.于是正方形被分为四个小正方形 1,2,3,4.由于大正方形的叠合关系,正方形 1 和 3 表示投影平面的同一多边形.同理,对于正方形 2 和 4 也是一样.其次,线段 PM 和 QM 必须看成一个棱,因为 P 和 Q 表示同一个点.同样,RM 和 SM 也应该看做一个棱,因为 R 和 S 表示同一个点.顶点只有一个,即 M.因此在欧拉公式里须令

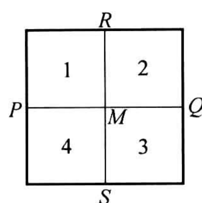


图 21.31

$$E = 1, K = 2, F = 2$$

代入公式后,得 $E - K + F = 1 = 3 - h$,从而得投影平面的连通度是 2,这正是表 21.1 中所列出的.

在解析投影几何里,用三线坐标可得投影平面的另一种划分法.这种划分法如图 21.32 所示,其中当做投影平面的模型的不是正方形面而是圆面,不通过同一点的三个圆弧划分圆面为七个区域.假定每个圆弧交圆周于二对径点,从而区域 2 和 5 表示一个三角形,3 和 6,4 和 7 也各表示一个三角形.可以看到,不通过同一点的任何三条直线必如此划分投影平面为四个三角形.现在将 $E = 3, K = 6, F = 4$ 代入欧拉公式里,也得到 $h = 2$.

我们曾经把正方形模型的对边黏合而得环面和克莱因瓶,应用同样的方法,也可以由正方形模型得到投影平面.为此先把正方形变形使成球面,挖去一小块四边形 $ABCD$(图 21.33).再把 AB 和 CD 黏合,DA 和 BC 黏合.这一步可以这样做:提高 A 和 C,拉下 B 和 D,然后将这二对点分别合在一起(图 21.34).结果就得出自交于一直线的闭曲面(图 21.35).这曲面跟投影平面拓扑等价.

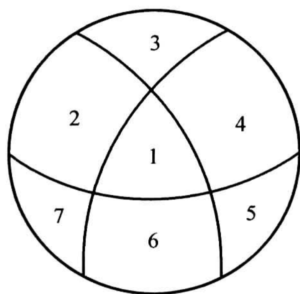


图 21.32

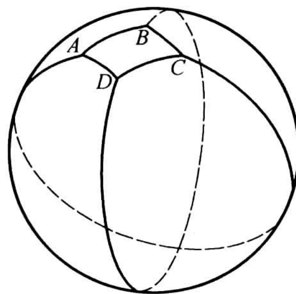


图 21.33

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

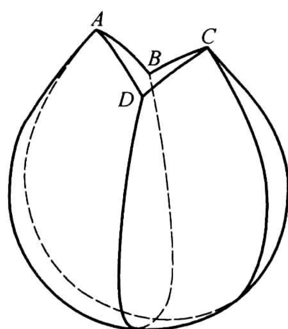


图 21.34

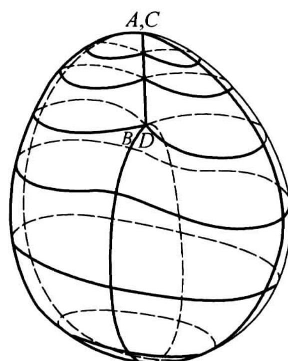


图 21.35

的确存在具有这种形状的代数曲面(图 21.36),其方程为

$$(k_1x^2 + k_2y^2)(x^2 + y^2 + z^2) - 2z(x^2 + y^2) = 0$$

这个曲面同微分几何中一个作图有关.列方程系对于直角坐标系而言,此坐标系以 P 为原点,以曲面 F 在点 P 处的二主方向为 x 轴和 y 轴,k_1 和 k_2 则为 F 在点 P 的主曲率.

如果我们再从图 21.34 出发,但现在只把 AB 和 CD 黏合而不把 DA 和 BC 黏合,便得到跟麦比乌斯带拓扑等价的曲面.因为这种叠合对边的方式,根据定义,正是变换正方形为麦比乌斯带的叠合方式.新曲面的边缘系从弧 DA 和 BC 得来.但因为 A 和 C 重合,B 和 D 重合,所以这二弧做成一闭曲线,这曲线的形状,可以 —— 比方说 —— 作成圆形(图 21.36).显然这个曲面没有自交线.在 A, C 和 B, D 重合的二点处,曲面的曲率不连续.但若把这两点的邻近加以变形,则得一到处连续弯曲的曲面.图 21.37 表示这种曲面.

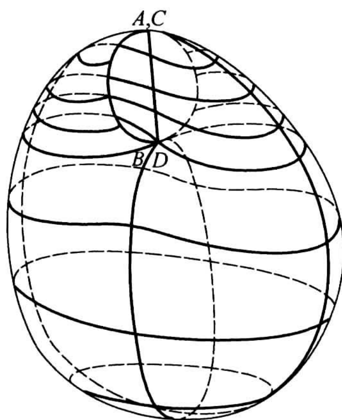


图 21.36

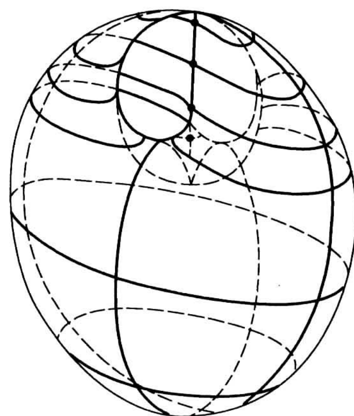


图 21.37

这样的曲面，尽管边缘是圆形的，却不能拿它作容器用，因为它是单侧的，没有内外部之分.

假若插入一圆面将曲面封闭，便重新得到投影平面的模型，这一点不难从图 21.36 和 21.35 上看出来.反过来，如果从投影平面的模型挖掉一片圆面，结果得一麦比乌斯带的模型.挖开的洞不管在图 21.35 的曲面上什么地方都可以，因为球面上各对对径点是类似的，投影平面上各点的位置也就没有区别.特别，如从图 21.35 的曲面去掉底部，剩下一个特别显眼的形象.这曲面名为"交叉帽"，如图 21.38 所示.交叉帽是以圆周为边缘的一种新的麦比乌斯带的模型.虽然曲面是单侧的，可是显然能作容器的盖子用.其所以可能，是因为它有自交线.

假如将变叉帽沿自交线剪开，再作适当的变形，结果得一有四边形孔或圆形孔的圆面：这只要把图 21.33 ~ 21.35 上的变换逆推即得.因此我们可从介于二同心圆的平面域出发，把小圆的每对对径点黏合起来而得出麦比乌斯带的模型(图 21.39).最初看来，似乎认不出这图形所代表的曲面就是在表 21.1 上正方形模型所代表的曲面.但是正方形模型确实可以从图 21.39 得出：把平环面剪成两半(图 21.40) 成两半，加以变形(图 21.41)，翻转一半(即对调图 21.41 上 e 和 b' 的位置)，最后把原来在一起的一部分棱实在地黏合，另一部分则抽象地作成对应(图 21.42).

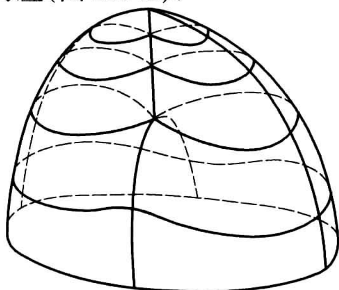


图 21.38

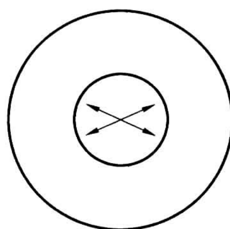


图 21.39

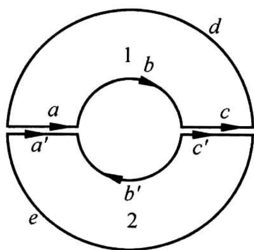


图 21.40

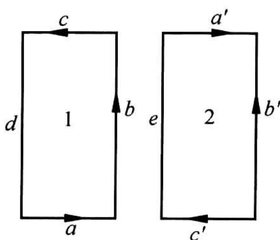


图 21.41

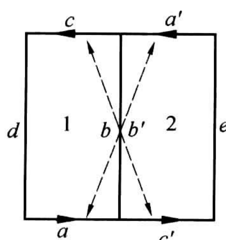


图 21.42

在我们的投影平面的模型里，有两点 —— 自交线段的端点 —— 是奇点. 鲍埃(W. Boy) 作成了投影平面的既无奇点又处处连续弯曲的模型.

为了作鲍埃曲面，我们不从正方形开始，而从正六边形开始. 六边形边缘上的对径点仍旧当做是叠合的. 把六边形变形使成球面，球面上有一球面正六边形孔. 新图形跟原来的六边形一样，可剖分为对于一轴对称(图 21.43) 的三片合同图形. 然后，把三片分离，每片再经一次变形. 至于如何变形，等一下再讲. 如此我们得到三片合同的成某种形状的曲面，把它们拼在一起，最后得到鲍埃曲面. 这曲面也有三重对称轴. 自然，这一方法的目的就是设法叠合球面的六边形洞的棱上的对径点.

首先，把一片曲面 $ScAaBbCdS$(图 21.43) 的三点 A, B, C 重合在 N 处(图 21.44)，但不当做是叠合的，因为不符合原来的叠合规则. 现在保持点 S, N 和棱 b, c, d 不动，把闭棱 a 提高(图 21.45) 到如图 21.43 的位置. 在这一步手续中 c, a 间的部分曲面自然要拉长一些，使它几乎成为平面. 然后把弧圈 b(图 21.46) 向右上方拉，使它处于刚才所说的那部分曲面的下面，如图 21.47 上的位置. 这最后的位置应使弧 c, d 合同，弧圈 a, b 合同，并且当曲线绕轴 SN 作 $\frac{2\pi}{3}$ 角的旋转(如图中所示的指向) 后，c 重合于 d，b 重合于 a(图 21.47). 我们再作一片与第一片合同的模型，它的对应部分以 a', S' 代表. 把这片模型粘在第一片上，使 d' 和 c 重合(而且 S' 和 S 重合，N' 和 N 重合). 于是 a' 和 b 自然重合. 把这二棱粘在一起. 现在，这两片模型的接缝线 Nb 变为新曲面的自交线，从图 21.47 上可以看出. 新曲面的边缘由 c', a, b' 和 d 组成，这只要回头看到原来的正方形即可知道(图 21.48). 显然，我们还可在这条边缘上粘上第三片模型，使 d 和 c'' 重合，a 和 b'' 重合，b' 和 a'' 重合，c' 和 d'' 重合(记号的意义同前). 至此，鲍埃曲面塑造成了. 从图 21.48 上容易看出，鲍埃曲面跟投影平面等价. 图21.49 系用金属丝编织的网状模型. 曲面的自交线是有三个弧圈的曲线，它通过点 N，而且同整个曲面一样，对称于 SN 轴. 仔细分析图 21.48，可知有三片曲面通过点 N. 这三片曲面在点 N 有连续切面的必需和充分的条件是在弧圈的六个端点汇聚的 N 处有三条相互垂直的切线. 如果曲面在其他接缝的地方有皱纹或曲率不连续的情形，可把曲面弄下来除掉. 在图 21.40 上，模型的自交线以粗金属丝做成. 至于其他的粗金属丝只是为了作支架之用. 这个模型用一个螺丝钉钉在点 S. 模型与我们的做法之间的关系，看图便十分了然.



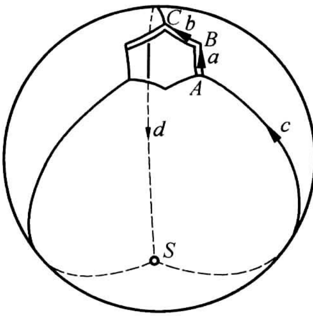


图 21.43

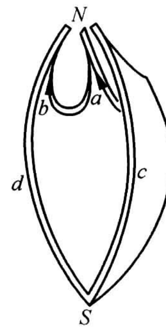


图 21.44

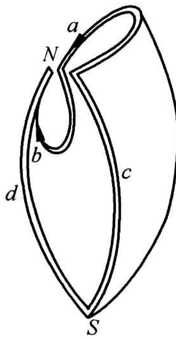


图 21.45

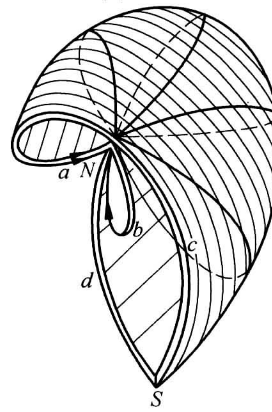


图 21.46

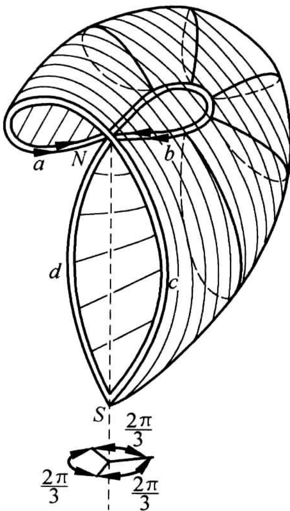


图 21.47

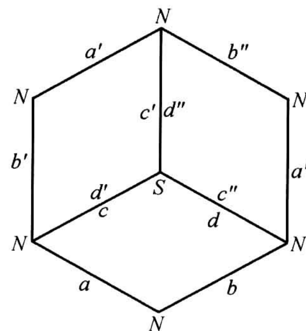


图 21.48

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

根据前面的讨论,可知鲍埃曲面的球面象处处连续.可惜现在还没有弄清楚球面象是怎样分布的.假定我们从任意法矢量出发,连续作出象点,则由于曲面是单侧的,终于会达到这点的反方向法线矢量.这样,在球面映射之下,鲍埃曲面上任何一点对应球面上的一对对径点.但由于在这样的叠合方式之下球面重新变回投影平面,可见鲍埃曲面的球面表示产生投影平面到自身的映射.当然映射不是相互单值的,因为显然鲍埃曲面上几个点对应球面上一对点.

5 有限连通度曲面的标准形式

我们把所有能互相拓扑映射的曲面划为一类.具有有限连通度的属于同一类的二曲面,其必需条件为:

(1) 二曲面或者都是闭合的,或者有同样数目的边缘曲线.

(2) 二曲面或者都可定向,或者都不可定向.

(3) 二曲面的连通度相等.

第一个条件的必需性很明显.第二个条件的必需性可以这样表述:凡可以拓扑映射成可定向曲面 G 的曲面 F,它本身也必可定向.在这种说法下,命题容易证明.事实上,曲面 G 的定向经过拓扑映射而变为 F 的定向.同样容易推证第三条件的必需性:曲面的连通度决定曲面的截口组的存在,而在拓扑映射之下,曲面上的截口组和曲面象的截口组的构造是一样的.

进一步研究表明,上述三个条件对于保证二曲面可以相互拓扑映射也是充分的.因为如果已知给定的曲面为可定向或不可定向的,又知其边缘曲线数和连通度,那么就可以仿照环面和连通度为5和7的可定向闭曲面的办法(参看图21.13 ~ 21.18),沿适当的截口组把曲面切开.切开后曲面总可以变为有一部或全部棱是叠合的多边形.截口组的机构,以及多边形的棱数和棱的叠合方式,完全取决于上述三个条件.因此二曲面如果同时符合三个条件,则必能经过拓扑映射而变成同一多边形域,因而能相互拓扑映射.

按照以上方式剖分的结果,亏数为 p 的可定向闭曲面变为 $4p$ 边的多边形,其叠合边配对的情形如图21.49所示.这些 $4p$ 边多边形作成代表所有可定向闭曲面的一个标准形式组,因为任何这类曲面的连通度是奇数 $h = 2p + 1$.我们见过的另外一个完全标准形式组是球面、环面以及有 p 个洞的圈饼.

在函数论中有几种黎曼曲面包括在这种分类之内,虽然从直观形象看不出这种可能性.这些曲面,类似大多数极小曲面的球面象,把球面覆盖了好几层,并且各层在支点处互相连通.这些曲面都是可定向的,因为球面的定向可变为覆盖在球面上黎曼曲面的定向.当且仅当函数是代数函数时,对应的黎曼曲面是闭合的;所有超越函数都产生非闭黎曼曲面.我们不再细讲,有许多关于几何



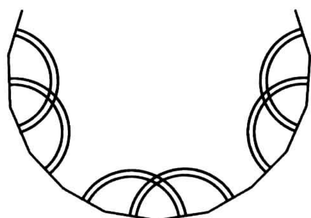


图 21.49

函数论的好书可供参考.

对于有边缘曲面,也可能找出一组多边形,使得边缘由有限条曲线组成的而且具有有限连通度的任一曲面,都恰可拓扑地映成这些多边形之一.圆筒面和麦比乌斯带的正方形模型是这样的多边形的两个例子.对于可定向的有边缘曲面,还可把球面、环面和圈饼随意挖上几个洞(图21.9),得到更为直观的标准形式.为了对于不可定向曲面作出类似的标准形式,可以利用交叉帽.在一球面上开若干个洞,然后用交叉帽把其中的几个洞封闭起来.任一具有有限连通度的不可定向的曲面都等价于用这种方式作出的曲面.交叉帽的数目和洞的数目唯一决定于边缘曲线的数目和连通度.

交叉帽有一条自交线和两个奇点.克莱因瓶和鲍埃曲面则提供了无奇点的单侧曲面的例子.关于其他一切不可定向的闭曲面是否都能以无奇点形式在空间里实现的问题,似乎尚未解决①.当然这类曲面绝不能没有自交线,这一点我们早已提过了.

然而,在四维空间内,所有不可定向曲面都可以表作既无奇点又不自交的曲面.今用 R_4 代表四维空间,用 R_3 代表三维空间.我们可以想象 R_3 安装在 R_4 中就像平面安装在 R_3 中一样.现在先在 R_4 中作一个无自交线和奇点的交叉帽.为此目的,设想在 R_3 中有一交叉帽,并设想它安装在 R_4 中.然后在交叉帽上选择一个以自交线为直径的圆面 e(参考图 21.36).在 R_3 中,我们可让任何圆面的圆周固定,而将圆面的其他部分弯曲,使所有点都离开圆面的平面.同样,在 R_4 中可把圆面 e 变形成为曲面 f,使圆周保持在 R_3 中的交叉帽上不动,而 f 的内部伸出于 R_3 之外.经过这次变形,交叉帽变为 R_4 中的曲面 F,而 F 显然是无自交线和奇点的.如果现在我们把有若干个洞的球安装在 R_4 中,不用交叉帽而用像 F 那样的曲面来盖住几个洞口,那么就得到一组既无自交线又无奇点,代表所有有限连通度不可定向曲面的标准形式.

另外一个问题是用最低次数的代数方程表达具有已知拓扑结构的曲面.例

① 俄译本将这句话改为"其他一切不可定向的闭曲面都能以无奇点的形式实现,如果对于克莱因曲面或鲍埃曲面安上适当数目的手柄的话".

如,我们从前讲过施泰纳曲面是代表投影平面的模型,那么有没有代表鲍埃曲面的代数曲面呢?这个问题尚未解决.投影平面可以实现为 R_4 中的一曲面,曲面的方程很简单,而且无自交线和奇点.

拓扑等价问题已经从曲面推广到三维和更高维的结构.这问题引起贝蒂(Betti)群的研究.在贝蒂群论里,曲面连通度和可定向性的概念是从更一般的观点来处理的.

6 将曲面映成自身的拓扑映射,不动点,映射类,环面的汎覆盖曲面

将曲面映成自身的最简单的拓扑映射是一个连续畸变,它把曲面整个变成自身,这类的映射叫做形变.平面到自身的运动是形变,但平面对于一直线的反射则是映射而不是形变.因为,反射反转了每个圆的指向,而形变则不然.

在映射之下映成自身的点叫做映射的不动点.现在我们证明,将圆面(包括圆周上的点)映成自身的任何连续映射,至少有一不动点.用反证法证明.假定圆面 e 上有一个将它映成自身但无不动点的连续映射.那么我们可给 e 中的每一点 P 加上一个箭头,使箭头的方向指着象的方向(这个规则仅对于不动点才不适用).因为由假设,映射是连续的,从一点到另一点箭头方向的改变一定也是连续的.现在考虑在圆周上一点的箭头,并让这点以顺时针方向绕圆周一圈;显然这点上圆的切线同时以顺时针方向转了整整一圈.我们现在断言,这点的箭头方向也以顺时针方向恰好转了一圈.首先,箭头转过的圈数必然是整数(也可以是零),因为它要回到原位置.因为圆周上一点的箭头永远指向圆的内部,切线和箭头的夹角不能等于零,也不等于 π 的任何倍数.但是如果箭头转过的圈数和切线转过的圈数不同,那么在圆周上至少有一次应该碰到或者相同或者相反的两个方向.其次,用同样的方式我们考虑和上述圆同心且在其内部的任一同心圆 k 上箭头所转过的圈数.这里也是一样,当一点以顺时针方向绕圆周一周时,它的箭头也以顺时针方向恰好转了一圈.假如不然,则在外圆周连续地收缩到 k 的位置的过程中箭头转动的圈数至少有一次起不连续的变化,这就跟圆内箭头分布的连续性矛盾.另一方面,如果 k 连续地收缩到圆面的圆心 M 处,则在 k 上所有点的箭头的方向必将趋近于一个方向,即点 M 的箭头方向.因此对于充分小的圆来说,圈数应该等于零.这就引出了矛盾.所以将圆面映成自身的连续映射不能没有不动点.

同样,我们可以证明,将球面映成本身的任一连续映射必有一不动点,或者有一映成对径点的一点.假如不然,每一点将唯一地决定连接这点及其象点的大圆的弧.这样势将得出球面上的一个箭头场,它是处处连续的,可是若考虑到





箭头的旋转圈数,可以证明这样的场是不可能的.由此可知,我们不能够在地球上竖立指路牌,使指路牌方向处处连续变化.

如果把对径点叠合成一点的球面解释为投影平面的模型,我们得到——作为上述球面映射定理的推论——将投影平面映成本身的任一连续映射必有一不动点.

要想更好地了解将已给曲面映成本身的拓扑映射的全貌,可以把所有的这些映射分类.如果二映射只相差一个形变,把它们列入同一类;诸形变本身作成恒等类.把球面上任一点映成它的对径点,就得到不属于这类的映射的例子;因为在这样的映射之下,小圆的指向显然反转了.如此,到现在为止,我们知道球面映射有两类.进一步研究可以证明(但离题太远),在球面上再没有别的映射类.由此可知,投影平面的所有拓扑映射都是形变.

另一方面,在环面上有无穷多类.为了考察其中几个类,设想把环面沿经线切开,拉直使成柱面,其二边缘为圆.然后将一圆固定,把圆柱面扭转,使圆柱面变成自己,而另一圆绕其轴旋转 k 圈;这样就将柱面的每一母线变成绕柱面的轴旋转了 k 圈的螺旋线.今将柱面再弯曲过来使成环面,并将二边缘圆粘在一起,结果是一个将环面映成自身的拓扑映射,同时边缘圆上所有叠合起来的点是不动点.其他诸点的映射决定于柱面的映射:柱面的母线对应环面的纬圆;此外,若是将二曲面的对应关系扩充到位于它们内部的部分空间,则可使柱面的轴对应环面的"心轴"(心轴即环面母圆中心描出的轨线).如此作出的环面在本身上的映射将纬圆变为环面上绕心轴 k 圈的螺旋状闭曲线.不论以后将环面怎样变形,都不能改变与这种曲线相伴随的数 k.因此环面上面两个有不同的 k 值的映射不能属于同一类.

假如认为由此可以类推,说克莱因瓶也有无穷多的映射,那就错了.在圆柱面上的螺旋线对应在克莱因瓶上的闭曲线,这些曲线,即使 k 的值不同,也可以相互变形.要问何以克莱因瓶和环面有如此差别,这可借正方形模型来说明.事实上,克莱因瓶仅有有限个映射类.

上述方法未能包罗所有的环面映射类.我们可以借助环面的汎覆盖曲面,以概括这些映射类的全貌.为此,我们设想把欧几里得平面包裹无限长的圆柱面;自然,这时圆柱面被平面包裹了无穷多次.从前我们不止一次地把有二边缘圆周的有限长柱面弯曲成环面.同样,无限长的柱面也能够变成环面.在弯曲的过程中,柱面的轴要绕环面的心轴无穷多圈,而柱面则被套进它自己无穷多次.用这种方法,能够把欧氏平面拓扑映射成为一个覆盖环面无穷多层的无皱痕无支点的曲面.这种曲面就是环面的汎覆盖曲面.

每绕环面的经圆或纬圆一圈,覆盖曲面从一层进到另一层.我们在环面上作典型截口组(包含一经圆和一纬圆),而以通常的方法将环面变作以一定方式

叠合了对边的矩形面. 现在如果在覆盖曲面上标出所有在典型截口组上的点, 然后将覆盖曲面变回成平面, 那么标出的点将形成一直线组, 它们匀划分平面成无穷多的矩形, 其排列的情形正像平面平移结晶群的基本区域. 每一矩形相当于覆盖曲面的一层. 为了弄清楚这一点, 让我们换一种方法来作汎覆盖曲面. 今以对边有一定叠合方式的正方形代表环面. 仿照平面正方形点格的做法, 把这些正方形拼在一起作成长条 S, 以二平行直线 a, b 为边缘, 向两方无限延伸. 若将 S 适当地弯曲, 使 a, b 重合, 则 S 变为无限长的圆柱 C. 组成长条 S 的若干正方形划分 C 为以圆周为边缘的许多区域. 将组成边缘的二圆周叠合, 可从这样的一个区域又变回环面. 如果采取像上面所讲的方式, 把柱面覆盖在环面上, 则所有区域将相互重叠, 而且每一区域恰好覆盖整个环面一次, 其全部边缘线重合于环面的典型截口组. 现在我们也像作正方形点格一样地继续作下去: 把整个平面用 S' 那样的相邻长条盖满. 如果把平面包裹 C 无穷多次, 使 S 又变为 C, 则盖满平面的所有长条显然重叠在 S 之上, 而所有长条的正方形剖分和 S 的剖分重合. 现在把 C 再覆盖在环面上, 则平面上的诸正方形相互重叠, 而且它们的边缘落在环面的典型截口组上, 这正是我们要证明的.

第二种做法指出可用一种特别简单的映射把环面的汎覆盖曲面 U 变成平面 E. 因为如果把 U 上覆盖环面同一点的那些点称为等价点, 则 U 上的任何等价点组都能用 E 上的正方形点格代表. 将 U 映成自身的拓扑映射中那些将任一点变为一等价点的映射组成一个群, 我们把它叫做环面的基本群 (f). 如此说来, $U \to E$ 显然将 (f) 变换为变正方形点格成自身的平移群.

设 g 是将 U 映成自身的另一拓扑映射, 对这映射虽不要求一定把任一点映成它的等价点, 但要求它永远把等价点偶映成等价点偶. 于是 g 对应将环面映成自身的一定的拓扑映射 h. 为了看清这一点, 我们注意环面的任一点 P 被曲面 U 上某一无穷等价点组 (Q) 盖住. 但根据 g 的定义, 点组 (Q) 的所有的象 (Q') 盖住环面的单一的点 P', 因此 g 决定将环面映成本身的拓扑映射 $P \to P'$, 它可用 h 来代表. 反之我们也可以证明, 对于环面的任一已知映射 h, 能够在覆盖曲面上找到同 h 的关系有如上述的映射 g. 这样 g 只能在群 (f) 中的任一映射的限度内被决定.

应用上面的结果, 我们能够得到将环面映成自身的映射类的全貌. 这里只叙述结果而不加证明. 设每一映射 g 被 E 上的映射 γ 所代替, 而 γ 是 g 经映射 $U \to E$ 得到的. 设 $ABCD$ 为和 (f) 对应的在 E 上的平移群 (t) 的正方形基本区域. 设 $A'B'C'D'$ 为在映射 γ 之下 $ABCD$ 的象; 则平行四边形 $A'B'C'D'$ 必然是 (t) 的一个基本区域. 于是环面的映射 h 是一形变的必需和充分条件是, $ABCD$ 可经平移而重合于 $A'B'C'D'$. 环面的其他映射类对应于点格的一母平行四边形所





能取的其他形式,以及将正方形 $ABCD$ 映成自身的旋转与反射①.

汛覆盖曲面的概念,对一切曲面均可确定.对于可定向闭曲面,我们这样得到汛覆盖曲面:把一些 $4p$ 边形连接起来,并在它们之间建立一种对应关系,其方式和在环面的情形对正方形的做法差不多.不过,当 $p > 1$,基本群不能以欧氏平移群代表;但是能够以双曲平移群代表,并以 $4p$ 边形为基本区域(对于 $p = 2$ 的情形).对于有边缘的曲面,需要引入具有开的基本区域的欧氏平移群或双曲平移群.对于不可定向曲面为了基本群在度量几何中能实现,必须在欧氏和双曲平移之外再补充以欧氏和双曲滑动 - 反射.

7 环面的保角映射

有这样的问题:一个曲面能否保角地映成自身或者映成另一曲面?如果能够的话,究竟有多少种映射方式?当时我们限于研究和一圆内部或一圆面(包括周界)或一球面等价的曲面.汛覆盖曲面的概念使我们能够对所有其他曲面讨论类似的问题.这里只讲如何去找将环面映成本身或映成别的环面的所有的保形映射,因为对环面所用的方法也适用于其他曲面,而且环面的情形比较容易想象.在这里和以后,我们所讲的环面不仅是指一圆圈绕在这圆所在平面上且不与圆相交的轴旋转出来的回转曲面,也指与环面拓扑等价的任一曲面.

有一个"非此即彼"的定理,它是说任一曲面,凡是同圆的内部等价,也就是同欧氏平面等价的,能够保形映射成双曲平面或者欧氏平面.我们可以把这个定理应用于环面 T 的汛覆盖曲面 U,因为这一曲面适合定理的条件.因此,设 U 保形映射成平面 E,让我们首先提出这个问题:E 是欧氏平面还是双曲平面?

现在,基本群(f) 必是将 U 映成自身的保形映射群,因为在(f) 的映射之下,U 的任一区域的象甚至和那个区域合同.因此在保形映射 $U \to E$ 之下,群(f) 变换成将 E 保形映成自身的一个群(t).但是所有将平面映成自身的保形变换为已知:若 E 是双曲平面,它们是双曲刚体运动;若 E 是欧氏平面,它们是欧氏刚体运动和相似变换.此外,我们知道群(t) 和欧氏平面的结晶平移群有某种亲属关系:(t) 中所有的映射,除恒等变换外都没有不动点,而且群的基本区域是四边形.假如 E 是双曲平面,(t) 必是具有有限基本区域的不连续的双曲平移群.而且设法使读者相信,这样的群的基本区域至少有八边,因此 E 非是欧氏平面不可.可以用初等方法证明,任一不是刚体运动的欧氏相似变换有一个不动点.这样说来,群(t) 除了包含恒等变换外,只能包含没有不动点的刚

① 如果点格定义为在笛卡儿坐标下整数点的集合,而且如果 A' 经平移已与原点重合,则平行四边形 $A'B'C'D'$ 由点 B' 的坐标 a,b 和 C' 点的坐标 c,d 来决定.为了刻画环面的全部映射类,必须赋予 a,b,c,d 以适合条件 $ad - bc = +1$ 或 -1 的所有整数值.

体运动,这是说,它是平移.又因为(t)是不连续的而且有一有限基本区域,所以(t)一定是一种结晶平移群.

现在应用同样的论证于任一其他的环面T':设U'为T'的汎覆盖曲面;设保形映射$U' \to E$将T'的基本群变为E的结晶平移群(t').我们曾经讲过将环面变为自身的任一映射能够扩张到汎覆盖曲面.同样,任一保角映射$T \to T'$能够决定保角映射$U \to U'$,使得U和U'的对应点永远覆盖T和T'上的对应点.映射$U \to E$和$U' \to E$变换$U \to U'$为将E映成自身的保形映射a,a必然是欧氏刚体运动或相似变换.而且,a必然把平移群(t)变为(t').

这样我们证明了,将T映成T'的保形映射仅当群(t)可以经过刚体运动或相似变换成为(t')时才行.这个条件可用几何形式表出:令t_1是群(t)的最短平移.令t_2是(t)的不平行于t_1的平移中的最短平移.设m是t_2和t_1的长度之比$\frac{t_2}{t_1}$,因而$m \geqslant 1$.设a为这二平移所夹的角;为使a唯一决定,只要假定$0 < a \leqslant \frac{\pi}{2}$.设$m'$和$a'$为群$(t')$中的对应数值.则存在一相似变换把$(t)$变为$(t')$的充要条件是$m = m'$,$a = a'$(证明是初等的,留给读者).因此我们对于任一个环面T可对应以二数m和a,使T只能保形地映射到有此相同二数的那些环面上.这二数(或另一对和它们按某种一一对应关系联系起来的数)叫做环面的模数.

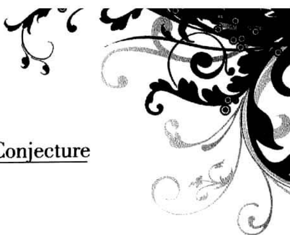
要使环面T和T'间存在保形映射,二模数相等不仅是必需的,也是充分的.因为如果条件满足,则在E上有一相似变换或一刚体运动a,变(t)为(t').容易看出,属于a的保形映射$U \to U'$决定一保形映射$T \to T'$,因为映射$U \to U'$变换U的等价点 —— 只是这些等价点而已 —— 为U'的等价点.总结上面的结果,我们可以说环面对于保形映射来说作成双参数族.

如果环面的空间形式没有什么特别的正则性,环面模数不能直接从它的几何形象上推出.但若环面T是回转曲面,则(t)一定有一矩形的基本区域,因此$\alpha = \frac{\pi}{2}$.因为,在这种情况下映射$U \to E$能够显明地给出:它将正交的经圆和纬圆网变为E上的二族正交直线.在特殊情况下,如果T是由圆产生的回转曲面,则t的矩形基本区域二边之比m仅依赖经圆半径和心轴半径之比.因此由圆产生的二环面能互为保形映射的充要条件是它们相似.

在四维空间里,我们可以找到一环面,其覆盖曲面U甚至能等距地映射成欧氏平面.

现在不难综合描述将一已知环面T映成自身的所有保形映射方法.这些映射所成的群(k)必然对应E上保持(t)不变的刚体运动和相似变换的群(l).显然(l)包含所有在E上的平移.一般地说,这些运动就构成(l)的全体.但如(t)





有特别的正则性,例如在正方形的基本区域,则(l)还包含旋转和反射.

上面对于环面讲的方法也适用于另外的曲面.但在大多数情况下覆盖曲面不像环面的情形那样保形地映射成欧氏平面,而是映射成双曲平面.这句断语对于例如所有亏数 $p > 1$ 的可定向闭曲面皆能适用.这类曲面引出双曲平移群,而且只有当二曲面的平移群可以经过双曲刚体运动而相互转变时,这二曲面才能相互保形映射.在双曲几何中可以证明,如果一个双曲平移群的有限基本区域有 $4p$ 个边,那么这个双曲平移群由 $6p-6$ 个常数决定.因此每一亏数 $p>1$ 的可定向闭曲面有 $6p-6$ 个模数.

在函数论中,这种方法主要地用在代数函数的黎曼曲面上.若 $p=1$,映射 $U \to E$ 导出椭圆函数;若 $p>1$,导出克莱因和庞加莱自守函数.

非闭曲面引出有无限基本区域的群.这种群时常在函数论中碰到,例如,在研究超越函数和椭圆模函数的时候.

神奇的二维国[①②]

第二十二章

1 关于这个国家

1.1 二维国的自然状况

这里所讲的世界叫“二维国”.这并不是我们随便给它起的名称,而是为使你们——有幸生活在三维空间的读者们——对它的本性更能清楚些才这样叫的.

设想平展着一张极大的纸,上面有直线、三角形、四边形、五边形、六边形以及其他图形;它们的位置不固定,能在纸面上自由移动,只是不能离开纸面做上下运动.它们很像是些影子,不同的是它们有明亮的轮廓,并且是硬的.这样一来,你该对我们这个国家及其国民们有一个较为明确的概念了吧!要知道,若是在几年前,我是会称它为“我的宇宙”的.可现在,我的眼界已经升级了.

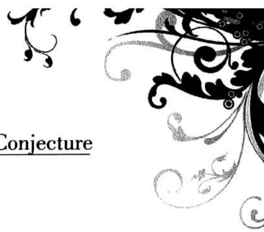
你立刻会觉察到,在这样的国家里不可能存在你们称之为“立体”的东西.可是,据我猜想,你一定认为我们至少可以靠视觉识别我刚才描述的那些可移动的图形,如三角形、四边形等.其实不然,我们看不见这样的图形,至少不能直接看出它们之间的差异.我们所能看见的只有直线,此外就既无所见,亦无

---

① 1884年,一位名叫艾勃特(Edwin. A. Abbott. 1838—1926)的英国牧师写了一本小书,书名叫《二维国》.这本书成了第一本构想种种不同维数的物质边界,并科学地讨论它们之间的关系的科普著作.

② 原载 Edwin A. Abbott. *FLAT LAND*. Dover publications. Inc 1952. 陈忱,译,北京:科学普及出版社,1991年.





可见了.我有必要马上来说明这一点.

把你们立体空间中的一枚硬币放在桌子中央,并从它的上方俯身向下看,你会看到它是圆的.然后你退到桌边,逐渐把眼睛放低(这便越来越接近二维国的国民所能看见的情况了),你就会发现这枚硬币在你的眼里变成越来越扁的椭圆形;最后,当你的眼睛与桌面完全在一个水平位置上时(此时,你的所见就同二维国的国民们完全一样了),这个椭圆就变成了一段直线.

你若用同样的方式看用纸片剪成的三角形或四边形等图形,其结果也会与上述情况一样.只要你的眼睛在桌边并与桌面平齐,它们看上去都不再是什么图形,而只是一段直线.以一个等边三角形为例 —— 它代表二维国中的一位处于可敬阶层的商人,图 22.1(a) 便表明你在他上方俯身向下时看到的样子,图 22.1(b) 代表你的眼睛近桌面的水平位置时看到的样子,图 22.1(c) 代表你的眼睛几乎与桌面平齐时所看到的样子.如果你的眼睛完全与桌面在同一水平线上,看见的便只是一条直线,正像我们在二维国里所看见的一样.

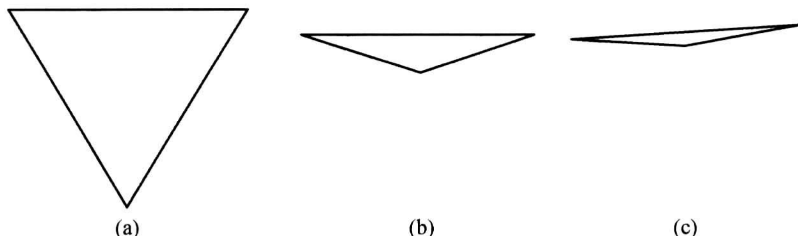


图 22.1

我在你们三维世界里曾听说,你们的水手也有相似的体验.当他们在海上航行时,会觉得天边的岛屿和海岸看上去也是这样.那遥远的陆地可能是左拐右折地蜿蜒着的海湾或地岬,可你在远处却看不到这些,只觉得水面上是一条灰暗的线在连绵不断地伸延着.(如果有阳光照耀,就会看到它是相间的亮暗线段的组合)

在二维国里,当这样一个三角形或其他老相识向我们走来时,我们所见的也正是这种情况.因为我们既没有太阳光,也没有别的什么光生成影子,因此没有什么能像你们在三维世界里那样帮我们观看.如果有人走近,我们会看见他的线段变长些,而在离去时则会交短些.无论是三角形、四边形、五边形、六边形,还是圆,我们看上去统统是一段直线.

你也许会问 —— 这是很自然的,在这种不利的情况下,我们如何识别朋友呢?随着我对二维国国民的描述,这个问题会很容易地得到合乎情理的回答.因此,我暂且将这个问题放一下,先把我国的风土人情介绍一二.

1.2　二维国的风土人情

同你们一样，我们的罗盘也指示四个方向：东、西、南、北.

我们这里没有太阳，也没有别的天体，用通常的方法来确定北方是不可能的.我们有自己的方法.根据我们那里的一条自然法则，二维国存在着一种恒定的向南的引力.它在温带处很微弱，因此连一般的妇女在那里也能毫不费力地向北走上几弗隆(弗隆是英国的长度单位，1 弗隆 $\approx$ 201.167 米 —— 原译者注)远.然而，在大部分地方，这种引力所产生的阻碍北行的效应也足以用来定向.此外，从北方送来的雨(降雨是定期发生的)也是一个附加的助力.在城里，我们的房屋可以用来指向，因为房屋的边墙多是修成南北向的，以使屋顶能遮挡北方降下的雨水；乡下没有什么房屋，我们就靠树干定向.总之，我们要确定方向并不像你们想象的那样困难.

在温带地区，向南的引力几乎感觉不到.有时，我走在那里的荒漠平原之中，找不到房屋和树木给我指向，结果不得不一连驻足几个小时，直等到下起雨来才继续我的旅程.对于体质虚弱或上了年纪的人，特别是对娇弱的女性，引力所起的作用比对强健的男子大得多.因此，如果男人在街上遇到一位妇女，总要让给她路的北侧 —— 这是个教养问题.能随时做到这点并非容易，因为男子们身体强壮，在温带地区很难辨别自己所处位置的北侧所在.

我们的房屋不开窗子，因为无论何时何地，无论白天还是夜晚，屋内屋外都是一样的.我们从来不知道光是从哪里来的.很久以前，学者们最感兴趣并经常研究的问题，就是光是如何产生的.但他们反复尝试研究解决的结果，只是使越来越多的人发疯.因此，在许多毫无结果的尝试之后，立法机构只好用课收重税的对策间接地阻止这种研究；近来，这一活动已被绝对禁止.在二维国里，目前只有我一个人完全知道这个神秘问题的答案.可是，在我们二维国里却压根儿没有人理解我，结果我反倒成了笑柄！我，一个掌握了三维空间的基本理论的人，却像是狂人中的最不堪者！还是不谈这些令人痛心的枝节东西，回到有关房屋的问题上来吧.

我们最普通的房屋是五边形结构，如图 22.2 所示.两条面北的边 RO，OF 构成屋顶，一般不在这里设门；东边有一扇专为妇女开的小门，西边是为男人开的一扇大门，南边是地板，通常也没有门.

出于下述原因，我们那里不许建造四边形和三角形房屋.因为四边形(这里是指正方形.后文中提及的各种多边形，除特别说明的之外，均应理解为正多边形 —— 原译者注) 房屋的墙角要比五边形的小，三角形的就更小些，而无生命物体(如房屋) 的边线比男人和妇女的边线更来得暗淡，所以，四边形或三角形房屋的尖角若被不经心的人撞上，就会造成严重伤害.按我们的纪元算来，早在



11世纪,法律便已规定,除了碉堡、兵营等民众未经允许不得接近的国防设施外,禁止建造三角形的建筑.

在这个时期,四边形的房屋虽不受禁止,但政府以加征特别税的办法予以限制.大约又过了3个世纪后,法律又规定,考虑到公众的安全,凡拥有一万人口以上的城市,五边形房屋的墙角是合法的最小构角.立法机关的努力得到了公众良知的反应.现在,即使在乡村,五边形的房屋也已取代了其他形式的建筑.只有在非常偏僻落后的农业区,考古学家才会发现一些四边形房屋.

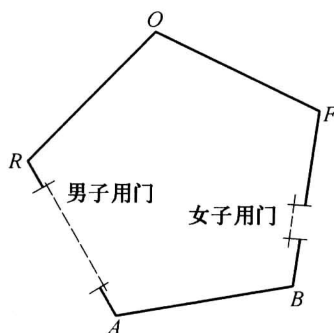


图 22.2

1.3 关于二维国的国民

二维国里最高的成年人用你们的尺度表示约为11英寸(1英寸 = 2.54厘米),12英寸被认为是我们人体高度的极限.

我们的妇女是直线.

我们的士兵和最下层的工人是有两个边相等的三角形.这两条边各长约11英寸,第三边也叫基边,它相当短(一般不超过半英寸),因此他们有一个极尖利的令人生畏的顶角.其实,他们的基边可以小到不足1/8英寸的程度,这时,他们几乎与呈直线状的妇女没有什么区别.为了区别于其他形状的图形,我们也同你们一样,把他们称为等腰三角形.以后我就这样叫他们.

我们的中产阶级由等边三角形组成.

我们的知识界人士是四边形(我本人便属于这一阶层),绅士则是五边形.

边数再多的就是贵族.他们又分为几个等级,从六边形起,然后边数逐渐增加,最高贵的被尊称为多边形.最后,当他们的边数增多(而边长同时缩短)到与圆没有多少区别时,就进入了圆的行列.圆是神职人员,属于最高的社会阶层.

我们还有一条自然法则:男孩总能比他父亲多生出一条边来.按此规律,每一代就会提高一个等级,直至成为贵族.这样,四边形的儿子是五边形,而五边形的儿子是六边形,以此类推.

但这一条规律不一定作用于商人,更少发生在士兵和工人阶层中.他们的边不全相等,这条法则不适用于他们.因此,一个等腰三角形的儿子依然是等腰三角形.然而所有的人(也包括等腰三角形)的后代都有希望最终上升到自己的等级之上.每当取得一系列军事胜利或经过勤奋的技术性劳动后,士兵和手

工业工人中的较聪明者的基边往往会有所加长，同时另外两条边有所缩短. 这些低阶层中的较聪明者，其儿女间的通婚（婚姻是由神职人员安排的）所产生的后代会更接近等边三角形.

在众多由等腰三角形父母产生的后代中，也会出现一些丝毫不爽的等边三角形. 但这种情况极少发生. 他们能取得等边三角形的证件①. 要造就出这样一个等边三角形后代，不仅需要他的先人进行多少代的精心安排的通婚，还需要他们进行长期不断的节制和自我控制训练，并一代又一代地耐心而有系统地持续提高自己的智力.

一旦等腰三角形的父母生出一个难得的纯正等边三角形来，周围远近的人都会将其看做一件大喜事. 经卫生与社会部门进行严格检查后，如果证明这个婴儿是规则的，就举行庄严的仪式接纳他为等边形阶层的成员，然后立即把他从自豪却又痛苦的父母身边领走，交给没有孩子的等边形收养. 养父母必须保证永远不让这孩子回他原来的家，甚至不让他把原来的家人看做亲属，以使他能健康成长，免得亲生父母的象庭影响潜移默化地使他再退回到先人的水平.

上述的偶然事件，不但使地位低下的下等人高兴 —— 这像一束微光，维系着他们摆脱悲惨境地的希望，也受到大多数贵族的欢迎，因为他们清楚地认识到，这种极少发生的现象非但不会削弱他们的特权，反而能使他们更有效地阻止自下而起的革命.

倘若有尖顶角的下等人无一例外地都绝无希望进入高级阶层，那么，他们就可能在暴动中拥立自己的领袖，他们会凭自己占上风的人数和武力，使大智大慧的圆阶层也难以应付. 然而，行事缜密的大自然做出了比喻：随着各劳动阶层在智力、知识和德行等方面的增长，他们那些令人生畏的尖角也会以相应的比率张大而接近较为安全的等边三角形的内角. 这样，最野蛮和最可怕的士兵阶层 —— 他们的智力几乎与妇女相当 —— 会在自己智能（这对他们决定如何使用自己的巨大刺穿力量是极有用的）的增长时，伤害力得到减弱.

大自然的互补法则是多么令人叹服！我简直要说，在我们二维国里，这种贵族社会构成的神妙起源是一个顺应自然的多么完善的证据！

多边形和圆们对这个自然法则的审慎利用，使人们对未来的无限希望不至破灭，从而几乎总能把暴乱扼杀在摇篮里. 技术也来给法律和秩序助一臂之力. 我们这里一般总能 —— 借助于医生略施医术进行人工牵引或压缩 —— 使一些较聪明的叛乱领袖变成完全规则的形状，并立即承认他们为特权阶层的成员；对于更多的不够标准的成员，则以最终会被封为贵族的诱惑把他们送入国立医

---

① 来自三维世界的批评家可能会问："要证件做什么呢?四边形的儿子自身，不就能证明他父亲是等边三角形吗?" 我的回答是：任何一位女士都不肯与一个没有规则三角形证件的男子结婚. 有时，一个稍不规则的三角形也会产生四边形的后代，但这类四边形的儿子几乎毫无例外地会表现出祖父辈的生理特点来. 结果非但不是五边形，还可能返祖为三角形呢.





院实行终身监禁;而对个别不肯低头的顽固分子则处以死刑.

至于那些乌合之众的等腰三角形策乱者,圆阶层的领导人则或以暗中收买的办法造成他们相互残杀,或更多地利用对方嫉妒和怀疑的弱点巧妙地进行煽动,挑起他们的火併而坐山观虎斗.在我们的历史上记载了一百二十多次大规模叛乱,还有二百三十五次规模较小的暴动,它们都是这样被平定的.

1.4 关于妇女们

如果说生有尖角顶角的士兵阶层是可怕的话,那么妇女简直就是可怖的了.因为若说士兵是楔子,那妇女就是针.之所以这样说,是因为她们的两端都是尖的.此外,她们还能随时使自己难以被人看到.因此,你可以想象得出,二维世界的女性是千万小觑不得的.

也许有些年轻读者会问:二维国的妇女怎么会让人看不见呢?我想,这个问题其实毋庸解释,原因是显而易见的.不过,考虑到可能有人仍不怎么明白,我就再说上几句吧.

把一根针放在桌上,使眼睛在与桌面平齐的位置上从侧面看它,你会完整地看见这根针;而若从一端去看它,则只能看见一个点,实际上几乎是看不见的.我们这里的妇女也正是这种情况.当一名妇女侧对着我们时,我们就能看见一段直线,而当她将生有一张嘴巴和一只眼睛 —— 这两个器官是生在一起的 —— 的一端转向我们时,我们便只能看见一个有光泽的小点了.一旦她将另一端对着我们,就连这点光泽也不存在了,简直就同无生命物体一样暗淡.因此,她们的后端就像是他们的隐身帽.(隐身帽是出现在欧洲许多国家童话里的宝物,谁戴上它,谁就会变得让人看不见.—— 原译者注)

妇女给我们带来的危险有多大,三维世界的读者现在一定清楚了吧!如果说撞上中产阶级可敬的三角形已经不无危险的话,撞上工人就足以造成划伤,与军官相撞则意味着受重伤,被列兵的尖角戳中就可能致死,而若碰到妇女的后端上,除了立刻呜呼哀哉之外怕不会有别的结果了.可是,当一个妇女是不可见的或只表现为一个无光泽的暗淡小点时,即使最谨慎的人也难免撞上哩!

为了把伤亡事故降到最少,二维国的各个地区在不同时期都颁布了许多法令.非温带区的引力较大,人们的运动较为困难,较为不自由,因而与妇女有关的法律自然也严格得多.下面三条法则概括了法典中有关妇女的方针:

(1) 每座房屋在东边应开一仅为女性使用的入口,所有女性都要"以得体而高雅的姿态"从此门出入①,而不得使用西边供男人出入的门.

(2) 在公共场合行走的女性必须不断地发出"轻柔的叫声",否则将被处

① 我在三维世界时了解到,你们那里的教堂也为村民、农夫和教师们设了专用门,并让他们"以得体而高雅的姿态"出入.

死.

(3) 所有患舞蹈病、痉挛或伴有剧烈喷嚏的慢性感冒等会引起不自主运动的女性病人,一旦正式体检查明病情属实便应立即处死.

在一些国家,对可怕的女性有一条附加的法律规定,即所有在公共场合行走或站立的女性须不停地摆动自己的身体,以使她后面的人知道其存在,否则也将被处死;还有的地方明令女人外出时须有儿子、仆人或丈夫陪伴;更有一些地方要求妇女除宗教节日期间外不得离家外出.但是,我们最贤达的国民和政界人士发现,对女性的种种限制如愈演愈烈的话,不仅有使种族衰微、人口减少的危险,也使国内的谋杀案激增.因而,太严酷的法律只能使一个国家失大于得.

被有关居家和外出的种种限制激怒了的妇女们,常会对他们的丈夫和孩子发泄愤懑.在非温带地区就发生过妇女同时行动,把全村的男人在一两个小时内统统杀光的惨案.上面提及的三条法规能符合管理较好的国家的立法需要,因此成了我们这里有关女性的法典的基本内容.

总之,安全环境的实现,不仅是立法的结果,也在于对妇女的切身利益有所考虑.因为一旦发生相撞时,虽然妇女向自身末端方向的急退能当场令人毙命,但除非她们自己能立刻抽身退出来,否则也会遭到自损躯体,甚至肢折骨裂的结果.

时尚的力量也于我们有利.我在前面说过,在文明程度不很高的二维地区,妇女们被明令在公共场合时必须不断摆动自己的身体.而据我所知,在所有治理良好的地区,追求大家风范的女子都养成了做这种动作的习惯.人们认为,无论哪一个国家,如果非要靠立法机关进行强制性实施才能维持的话,那是很不光彩的.上流社会的淑女是天生就具备这一懿德的.圆阶层的妇女都能够做出富于节奏感的、动作曼妙的摆晃,这使诸多等边三角形阶层为妇人们欣羡不已并竞相效仿,然而却只能像钟摆似地单调地晃动.不过,她们的这种有规律的动作,又使不少巴望着向上爬的等腰三角形的妻子们纷纷效颦,但总是连这一习惯也养不成.这样一来,在所有有地位的或者和睦的家庭中,妇女们便都有了摆动身体的习惯,从而使她们的丈夫和儿子至少不至于横遭无端伤害了.

这么一介绍,我们的妇女怕要被看做是缺乏感情的干面包了吧?其实不幸得很,脆弱的女性常为一时的激情所左右而置其他一切于不顾.当然,这是由她们自身不幸的结构必然引起的 —— 她们没有角,地位比最低等的等腰三角形还要微贱;她们智能低下,既缺乏思考判断,也没有远见卓识,还没有持久的记忆力.因此,一旦她们发起火来,就忘了自己的责任和身份.我确知这样的一件事:一个妇女毁了自己的全家,但半小时后,丈夫和孩子残躯刚刚被收拾干净,她的怒火也平息了,却又问起家里人都哪里去了.





显然,妇女若呆在有条件能转身的地方,男人就不应去激怒她们.当她们呆在自己的房间里时(女人的房间都设计得不让她们转身),你便可以自由自在地行事,她们全然无法干涉.由于她们的记忆力维系不了多久,因而尽管会因一点点小事就用死来威胁你,但过不了一会儿就时过境迁了,而且连你为平息她们的怒火而随便作出的许诺也忘得一干二净了.

除了军人这一较低阶层之外,我们的家庭关系一般是相当和睦的.军人由于不够稳重,不够明智,有时会捅出大娄子来.有些鲁莽的家伙不是把他们的尖角当做良好的防卫手段,而是频繁地将它作为进攻的武器;他们常常不按前述方式为女人修造房间,也总是在外面胡来而又不表示改悔.这些都会激怒自己的妻室.此外,他们又都认定一言既出,驷马难追这条死理,不像圆们那样会用甜言蜜语和漫天承诺来取悦太太.这些毛病都导致了残杀,然而,这也并非全无益处.许多特别野蛮和捣蛋的等腰三角形就这样被干掉了.我们的许多圆阶层成员还认为,细狭人等的减员是节制人口的最好方式,又可起到遏制革命的作用.

但是,即使我们那里最规则和最接近圆的家庭,在生活上也不能说有你们三维世界里那样和谐美满.没有残杀只可称做和睦,但在志趣和追求上却未必和谐.圆阶层的明达固然能保证家庭生活的安全,但其代价是舒适的消失.每一个圆或多边形的家庭中自古以来就有一个习惯 —— 现在则已成为我们较高阶层妇女的一种本能 —— 就是母亲和女儿总要把自己的眼睛和嘴巴一直对着丈夫和男友.倘若这些名门淑媛不这样做,就会被认为不吉利和失身份.你们很快会看到,这种习俗虽有利安全,可也有不利之处.

在工人或有身份的商人家中(这些人家的妻子在做家务时可以不把眼睛和嘴巴对着丈夫),当女人们除了只发出连续的轻柔叫声,而不让人看到她们在场时,至少会造成片刻的安宁,而这是较高阶层的家中无法实现的.高贵家庭中的妇女都伶牙俐齿、目光尖利,注意力又都放在一家之主身上.时光会在她们口若悬河的言谈中流逝,要机智巧妙地截断她们滔滔不绝的话头,怕要比躲开他们掉过身来的一刺还要困难.当做妻子的越来越没话找话,而做丈夫的越来越没有机谋、情绪或良知让她们住嘴时,真有不少男人宁可挨上她们那可怕然而无声的击刺而死,也不愿听着她们的喋喋不休而苟活呢!

三维世界的读者们会觉得,我们的妇女境遇实在悲惨.的确是这样.在我们这里,就连最卑微的等腰三角形男子,都有个人角度加大和整个阶层得到提高的指望.而妇女却没有改善状态的可能."一旦入女流,终生为妇人."这是一条自然的法规.正是这种进化规律在她们身上的冻结,造成了她们的不幸.不过,她们没有希望的命运,倒恰好同她们记忆力低下 —— 既记不住过去,也记不住被告知的未来 —— 的状态相匹配.因此,妇女和我们整个社会的基本构成都能

得以维系. 这不能不令人赞叹造物主的安排呀!

1.5　我们相互辨认的方法

在你们得天独厚的三维世界里,既有光又有影,你们每人还生有一双眼睛,懂得透视原理,并能享受各种色彩的魅力;你们能实际地看见一个角或一整个圆. 因此要让你们了解我们在二维世界里的相互识别的方法是十分困难的.

记得我在前面曾告诉过你,存在于二维国的所有物体,包括有生命的和无生命的,不论它们的形状如何,呈现在我们视野里的都是、或几乎都是清一色的直线. 那么,我们怎样互相辨认呢?

方法有三种. 第一种是靠听觉. 我们比你们有更发达的听觉,所以不仅能靠声音辨认出熟悉的亲朋好友,也能靠它识别对方的所属阶层,至少是等边三角形、四边形和五边形这三个最低的 —— 等腰三角形不算在内 —— 阶层. 但是,对于再高阶层的成员,就难由听觉进行辨认了,这一方面是由于这些阶层的发音彼此相近,另一方面是由于分辨声音的才能主要为平民具备,而在贵族中并不显著. 因此,凡在会出现冒名顶替的场合,就不全凭这种方法辨认. 最低阶层的发声器官比其听觉器官发达,这样一来,等腰三角形就很容易装出多边形的声音,如再经过一定的训练,还能装出圆阶层成员的声音来. 因此,我们更多的是使用第二种方法.

第二种方法是靠触摸. 这主要用在妇女和低等阶层中. (至于高等阶层,我马上就会讲到) 陌生人之间相互进行识别时 —— 不是识别个人,而是识别其所属的阶层 —— 用的就是触摸. 在你们三维世界,有教养的人初次见面时需要"引见一番". 我们这里则是要"引摸一番". 在远离城市偏僻地区,乡绅们今天仍采用一种古老的、也是我们二维世界习用的方式相识. 他们初次介绍两个人相识是这样说的:"请允许我请求您摸一摸我的朋文某某先生,也请您让他摸一摸." 不过在城市里,"让他摸一摸" 这一部分如今已被省略掉了,于是这句话便成了"请允许我请求您摸一摸我的朋友某某先生." 当然,话虽是这样说,"摸一摸" 仍是相互间的. 而在不愿多费一丁点儿事,也不在乎保护自己语言的纯洁性的现代派新潮青年中,这种惯用语句又被进一步缩短,于是便成了:"X 先生,请摸 Y 先生." 之类的话. 此时,"摸" 已成了专门术语,意为"允许我请求您摸一摸同时也被摸一摸".

然而,请读者们不要认为,触摸对于我们也像对你们一样令人不快,或者认为我们在判断别人属于哪个阶层时必须接触它的所有边. 从打上学时起,又经过实际生活的长期训练和实践,使我们用触觉能一下子判断出等边三角形、四边形和五边形的不同角度,至于等腰三角形那没头脑的尖顶,更是触觉迟钝的人也能摸得出来的. 一般来说,我们只要摸一个角,就可以判明这个人所属的阶





层,只有对贵族还做不到这一点.判断他们要困难得多.我们有一所著名的学府温布里奇大学,连那里的文科硕士也摸不出对方十边形抑或十二边形,至于究竟是二十边形还是二十四边形,那就连科学博士也未必能摸出来了.

凡还记得我在上文中提到的有关妇女法典的大致内容的读者,都会立即想到,进行触摸而结识人时应谨慎小心,否则触摸者便可能因不经意而受到对方尖角的不可补救的伤害.为了触摸者的安全,被触摸者一定要一动不动;无论是惊吓之下的一挪,还是毛毛躁躁的一动,喏,甚至只是打一个大喷嚏,都会产生致命的后果,从而葬送刚要开始的也许有希望长期维持的友情.对于较低阶层的三角形更是如此,他们的眼睛离尖顶是那么远,以至于他们几乎不能看到那里的情况,而且他们又天性粗犷,对讲究条理的多边形的轻轻触摸很不敏感.因此,他们也许只是偶然地晃上一下,就会夺去一条高贵的生命.

我的祖父是位顶呱呱的好人.他在所从属的不幸的等腰三角形阶层中不规则程度最低.他在作古前得到了卫生与社会部门的认可而成为等边三角形——这是以四比三的表决结果通过的.我常听他流着痛苦的泪水回忆本家族往昔的一桩过失.他的太祖父是一位受人尊敬的工人,顶角是59°30′.我这位不幸的先祖遭受着风湿病的折磨.一次,当一个五边形触摸他时,他悸动了一下,于是不幸斜向刺穿了这位大人.这给他带来了长期的牢狱之灾,而且又株连到所有亲属.他们的顶角都离他们争取的角度远了一度半.在他的曾祖父一代,登记的角度只有58°,直到第五代上才实现60°,从而从等腰三角形阶层提升到等边三角形阶层.而这一系列的苦难只是一次轻轻地触摸引起的.

至此,我仿佛听到了诸位受过良好教育的读者们在问:“在二维国里,你们怎么会了解度啦、分啦、秒啦的呢?我们处在三维空间里,所以能看见角,能看见相交的两条直线,或是位于一条直线上的若干小段.你们辨认不出角来吧,恐怕更分辨不出角的大小来吧!”

我的回答是:虽然我们看不见角,可我们能推断它们的存在,而且能十分精确.在需要的促进下,也在长期的训练下,即使不用尺和量角仪器,我们也能比你们用视觉更精确地辨认各种角度,此外也不应忽略另一点,就是我们还有“天助”——这就是我们的又一条自然法则:等腰三角形的顶角从半度即30′开始,每一代可以增长半度,一直增到60°止.到这个角度上,他们便可摆脱底层地位,成为自由人而步入规则图形的阶层.

可见,大自然为我们等腰三角形提供了顶角从30′增到60°的升级图,全国的每所小学校里也都有活人的陈列.由于偶然的退化,以及多过前者的道德和智力的停滞,再加上罪犯和流浪汉的高生育能力,因此社会上总有不少顶角为30′或1°的等腰三角形,顶角小于10°的更不在少数.他们没有公民权,其中很大一部分甚至连当兵都不够格.政府只得把他们提供给教育界消耗.为避免伤人,

需把他们锁绑起来，放到小学校的教室里，用以教育中层百姓的子弟，使学生们得到有关他们的知识. 而提供这些知识的可怜虫自己却连这点关于自身的常识都不了解.

有为数不多的学校给这些活标本东西吃，因此使他们能苟活几年，但在法制更强化的温带地区，不供食物而每个月 —— 在不进食的情况下，罪犯的平均存活期为一个月 —— 更换活标本的办法，从长远针对教育青少年有更好的效果. 在条件较差的学校，由于缺乏购置活标本的经费而让标本进食以多活些时间的作法，固然能省钱省事，但是这点好处，一部分会被提供食物的花费抵消，一部分会因标本连续好多个星期被学生们触摸造成磨损所导致的教学效果下降抵消. 频繁更换还有一条很重要的优越性，就是它有助于减少过剩的等腰三角形. 这种减少在效果上并不十分明显，但仍能觉察得到. 二维国的所有政治家一向都持这一观点. 我知道许多校董会欢迎少换标本的“廉价系统”，但我认为，真正的经济是消费，对标本的选用也不例外.

不能让有关学校的问题来转移我们的主题. 关于靠触摸来辨认的方式，我觉得自己已经说了不少了. 这种方法不像有人想象的那样费事、那样没有准头，触觉显然比听觉更可信赖. 但正如前所述，这种方式有一定的危险性，因此中低阶层中的一部分人，以及所有的多边形和圆宁可用第三种方法进行辨认. 这就是下一节的内容.

1.6 视觉辨认

这里又出现了视觉辨认，是不是同以前的说法不一致了呢？我在前面固然说过，在二维国里，各种图形统统表现为直线，但这并不意味着用视觉就不能辨认形体. 现在我就来向三维世界的批评家们说明我们是怎样用视觉来相互辨认的.

如果读者们再费心看一下前面所说的内容，就会发现那里说的是低等阶层中的情况. 在高等阶层中，以及在温带地区，人们是借助视觉实现辨认的. 各个地区和各种阶层的人们. 之所以能靠视觉进行识别，是因为有雾的缘故. 这种雾在一年中的大部分时间里遍布除热带以外的所有地区. 若在三维世界，雾对你们来说是不折不扣的坏事，它会遮蔽景观、压抑情绪，并且有碍健康. 可雾对于我们却不亚于空气那样重要；它堪称艺术之保姆，科学的母亲. 我还是别忙着唱颂歌，先讲清我的意思吧.

如果没有雾，我们看到的会只是清一色的长短直线，根本无法进一步区分 —— 在气候干燥大气澄澈的地方也的确是这种情况. 而在多雾地区，物体在一定距离外，比如 3 英尺（1 英尺 = 0.304 8 米）远处，看起来就会比 2 英尺处显得暗淡些. 因此，只要经常进行物体明暗程度的仔细观察，就能非常精确地辨认出一个个物体来.



一个例子会比一大堆泛泛之论讲得清楚明白.这里就通过例子解释一下.

设想有两个人向我走来,我想确知他们的身份.如果他俩一个是商人,另一个是医生;换句话说,一个是等边三角形,另一个是五边形,我将怎样辨认出他们来呢?

对于三维世界的孩子,只要学过一点儿几何知识,这个向题就是显而易见的.如果我的目光能等分开向我走来的陌生人所张的角度(A),我就正处在距他挨近我的两个边(即 CA 和AB)都不偏不倚的地方,我就能看到这两条边是等长的.

现在来看商人,如图 22.3(a).我能看见什么呢?能看到一直线 DAE,这条直线的中点(A) 离我最近,所以看起来很亮,但从点 A 到两端的两段线会随着它们很快隐入雾中而变得暗淡起来,我所能看到的三角形的两个端点,即 D 和 E,实际上就非常之模糊了.

另一方面,对于图 22.3(b) 所示的医生,虽然我也只看见一条直线($D'A'E'$),它的中点(A') 也是亮的,它的两个边($A'C'$ 和 $A'B'$) 也是逐渐变暗的,但它们向雾中隐入得不那么快,因此我看到医生的两个端点 D' 和 E' 也不如商人所对应的那两个端点那么模糊.

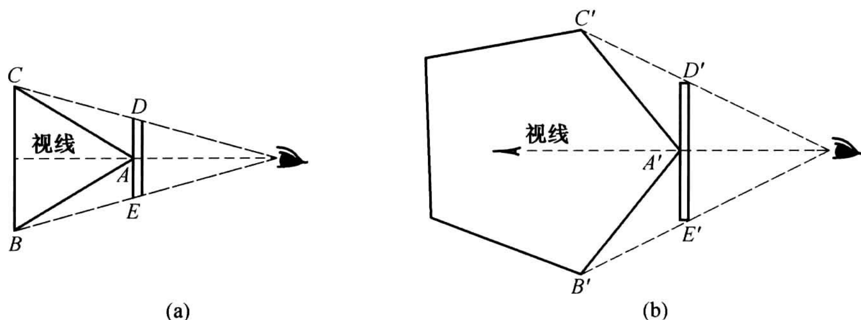


图 22.3

由这两个例子,读者们大概会明白,那些有知识的人凭借自己的学识,再加上长期的实际训练,是能够很准确地用视觉辨识中、低阶层的人们的.如果你们 —— 我的三维世界的读者们 —— 已经掌握了这个一般概念,至少就会认识到这样做的可能性,而不会认为这一切是不可思议的了.这也就达到了我讲这段话的目的.不过,如果让我进一步解释其中的细节,那我只得知难而退了.年轻而缺乏经验的人,可能会觉得 —— 根据前面我讲的两个简单的例子 —— 用上述方法也能容易地辨认出我的父亲和我的儿子,那我就必须指出,在实际生活中,大多数视觉辨认要比上述情况细致复杂得多.

例如,当我的父亲(一个三角形) 走近我时,很可能会以一条边而不是一个角正对着我,这时我就得让他转一转身,或者我自己围着他兜上几步路,才能把

他分辨出来，否则，此时我会怀疑我看到的是一段直线，或者说是一位妇女. 再有，当我的一个孙子(六边形) 陪着我时，他的一个边(AB) 会直对着我，从图22.4 中可以清楚地看出，在我看来他只是一条整个都比较明亮的，即使两端也不怎么变模糊的直线(AB)，再一边接上一段短短的暗淡的线段(CA 和 BD)，而后两条线段的端点 C 和 D 是较模糊的.

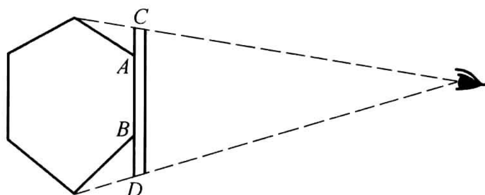


图 22.4

在我们这里，有教养的人在生活中有一个重要内容，就是运用他们在名牌大学里从著名教授那里学到的几何学的丰富知识 —— 有两种几何学，静止几何学和运动几何学 —— 用眼睛辨认出其他有地位者的角度. 由于辨认者自己是在运动的:或旋转，或前进，或后退，被辨认者可能人数众多，又都可能同时朝各个方向运动着，进行辨认实在是不容易. 对此，就是三维世界里最卓越的数学家也是会相信的. 但是我这里不准备再谈了.

只有极少数的最显赫最富有人家的子弟，才有时间和金钱全面学习这项高尚而重要的艺术，甚至连我这个名望不低的数学家，又是两个很有出息的完全规则的六边形的祖父，一旦置身在地位更高的一群转动的多边形中间，有时也会感到十分窘迫呢. 当然，普通的商人或者贫民，一旦置身于这种环境中，就会像你们 —— 尊敬的读者们 —— 一旦来到我的国家时一样的困惑和迷惘.

置身于这样的一群人之中，你会看到周围都是线，而且看上去都是直的，在这些线的各个部分上，明暗强度又都不断地无规则地变化着. 即使你在大学的五边形班和六边形班里学满了三年，出色地掌握了这方面的理论，也会感到还需要经过多年实践，才能在上流社会的人群中走动而不致撞上人，而这些人又把别人的触摸看成是十分无礼的举动. 由于他们有较高的文化修养，因此了解你的一切动作，而你对他们却了解甚少，或几乎是一无所知. 总之，要在多边形社会里应付裕如，自己就得是个多边形. 这至少是我亲身痛苦经历的总结.

说来也许令人惊奇，这种视觉辨认艺术有时简直可以说是成了本能，这主要靠反复实践和避免触摸培养成的. 这正如你们三维世界中的聋哑人，一旦学会了正规的哑语，就再也无法掌握凭口形识别正常人的语言这门更重要更困难的技术一样. 对于我们，“看” 就同这一情况一样，一旦在小时学了触摸法，“目识术” 就再也学不到家了.





因而,在高阶层人士中,触摸术是不受鼓励的,甚至干脆遭到禁止.他们的孩子自幼不是进公立小学读书(那里教授触摸术),而是被送入专门的高级学习班.在我们的最高学府里,触摸动作被视为严重的过失,初犯者受停学处分,再犯者予以除名.

但是,在那些低等阶层,视觉辨认术被看做是可望而不可即的东西.一个普通的商人是不可能拿出巨款让儿子花掉三分之一的时间去从事这种抽象问题的研究的.由于穷人家允许孩子从小对人进行触摸,因而在早年,他们若与多边形阶层的那些刚受到一半教育的孩子所表现的迟钝、不成熟和行事拘谨相比,就更显得成熟、活泼而讨人喜欢.可是当多边形的孩子最后完成了大学学业,并把他们学到的理论知识应用于实践时,他们表现出来的巨大变化却几乎可以说是一种新生;他们在文学、科学和社会学等各个方面,都会很快把三角形竞争对手远远抛在后面.

在大学里,会有极少数多边形子弟不能通过毕业考试,这少数落伍者确实可怜:上层社会会把他们拒之门外,低阶层的人也瞧不起他们.他们既没有多边形学士或文科硕士那种靠系统训练培养起来的成熟才干,也没有青年商人的那种天生的熟稔灵活和多才多艺.各种脑力劳动和国家公职也一概同他们无缘.虽然大部分地区并不明令禁止他们的婚姻权利,然而,他们也很难找到合适的伴侣.经验表明,那些不幸的无才华的父母所生育的后代可能会是不规则多边形,至少也是些不幸的人.历史上的骚乱和暴动,正是从这些出身贵族阶层的落难子弟中得到领袖的.这样的大错误使一些较进步的政治家提出,真正的仁慈总是完全的斩草除根.他们主张制定法令,把所有没有通过大学毕业考试的人,或者判处终身监禁,或者用安乐死的方式把他们从肉体上消灭.目前,这种人虽为数不多,但比例却正在上升.

我已经离开话题,谈到不规则图形上面去了.这是一个极为重要的问题,需要专门用一节的篇幅讲一讲.

1.7 不规则图形

前面我一直设定 —— 也许应当作为最基本的命题明确地放在开头挑明 —— 二维国里的每个人都是规则图形.由此,我说妇女是线时,就是说她们是直线;工人或士兵必须有两条边相等;商人一定要三条边都相等;而律师等 —— 我便属于这一阶层,但地位较低 —— 的四个边都是相等的 …… 总之,每个多边形的所有边都必须等长.

边的长短与年龄有关.刚刚出生的女婴长约1英寸,而一个成年的高个子妇女身高可达1英尺;各阶层的成年男人各边的总长为2英尺或略长一些.可这里并不考虑一个人的边有多长多短,重要的是边长应当相等.无须赘述便可看

出,整个二维国的社会生活都依赖于这样一个基本事实,即大自然赋予我们的所有形状都是等边的.

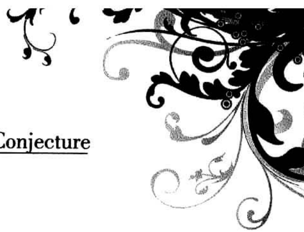
如果一个人的边长不相等,则其角度也不相等,这样,无论用触摸法识别还是用视觉法辨认,单凭一个角的大小是无法确定一个人属于何种等级的,只有通过触摸查明他的每个角才行.但是,这样费事地摸了又摸,实在是太浪费时间了.全部视觉辨认术也会立即告吹,就是触摸作为一门技术,也不会有长久的生命力.人们的交往将变成一种冒险.人们的一切打算将落空.信任将不复存在.即使是最简单的社交安排也不会有安全感.总而言之,社会文明将崩溃,人类将处于野蛮状态.

以上这些内容因为在我而言是显而易见的事,所以讲得可能太快了一点儿.不过,读者只要稍稍思考一下,我再举一个日常生活中的例子,大家就会相信,我们的整个社会体系都确实建立在人们的规则性上,也就是等角构造上.例如,你在街上遇到了两三个商人,只要看一眼他们的角度和亮度很快变暗的两条边,就能立刻认出他们的阶层,于是便会叫他们留步,邀请他们去你家用饭.你能非常自信地这样做,因为大家都知道一个成年三角形会占有多大的面积.可是你不妨想象一下,要是商人这个规则而体面的三角形尖顶后面又拖着一个对角线有十几英寸长的平行四边形,结果这个怪物卡在了你家门口,那该多么尴尬呀!

我怕是小看了有幸生活在三维世界里的读者们.我说的这些细节,读者们会是一清二楚的.显然,在这种怪异的情况下,只靠测量一个角已经不再能辨识别人了.所以,人们一生中需花许多时间来全面触摸他所要辨认的人的每一条边,即使是一个受过良好教育的四边形,要做到在人群中避免相撞也会是煞费周折的.可是,如果社会交往的成员不是规则形状的,那就会是彻底的混乱不堪;一旦有人稍一惊慌,就会造成重伤,若发生在妇女或士兵当中,则会是生命的巨大损失.

规则的人体结构是造物主赋予我们的天然标志和烙印,而自然法则也不会重新写过.不规则形对于我们,就如同你们所说的道德败坏和犯罪,而且这两者受到的处置方式也是相仿的.的确,我们这里不乏散布没有必要把几何形状的不规则和道德沦丧联系起来的人.他们发表评论说:不规则体自打一出生,就受到父母的歧视、家庭的忽视,以及社会的轻蔑与猜疑.他们与一切领导职务无缘,与一切受到重视的有用活动无缘,又时刻都受到警察的注意.成年之后,他们就得去有关部门报到接受检查,一旦发现他们的身体超过一定的偏差界限,就得送掉性命,在顶好的情况下也只是被雇用为最低级的雇员.他们不得结婚,干的是单调乏味的工作,得到的是微不足道的薪金,而且必须食于办公室、寝于办公室,甚至度假时也要受到密切监视.这对一个人的天性是多么大的摧残啊!





在这样的环境里，即使是最好最纯粹的人，也会产生怨恨而堕落的.

这一切听起来并非没有道理，但不能说服我. 明智的政治家们也同样不认为我们的祖先把对不规则图形的宽容与维护国家的安全不能兼顾是什么错误. 毋庸置疑，对不规则者本身来说，生活确实是艰难的，但大多数人的利益要求他必须这样. 如果一个人生着三角形的前身，而后背却像个多边形，又允许他生出一个更加不规则的后代来，那么生命的意义又何在呢？难道为了接纳这样的怪物，就去改建我们二维国的所有房屋、厅堂和教堂吗？难道戏院的收票员得先前前后后地测出每个人的各条边后，才允许他们入场吗？让不让形状不规则的人进入民团呢？再有，这种人也太容易受到冒名顶替的诱惑了呀！他们能把多边形的正面探进商店里购货，从而轻易地骗取商人的信任而捞取便宜. 让那些鼓吹废除对不规则形的刑法的所谓慈善家们去鼓噪吧，据我所知，不规则形天生就是些伪君子、厌世者，或者是些无所不用其极的恶棍. 大自然既能造出这种形体，也就能赋予这样的品性.

当然，我也不赞成一些地区采取的极端手段，如顶角与合格角度相差只有半度的婴儿一出生便得杀掉. 其实，在最高贵和最富才能，称得上是真正天才的人中就曾有一些是从角度偏差有 45′ 甚至更大些的逆境中奋争出头的. 如果当初消灭掉了他们，那真会是国家的不可弥补的损失哩. 我们在对不规则的治疗技术上也已取得了一些辉煌成就，可以采用牵拉、挤压、穿孔、诱导等种种外科疗法，把不规则者部分地或完全地矫正过来. 因此，我颇赞成中庸之道，不主张严格划定什么规则与不规则的界限标准，但是我建议，对不规则形应在形体开始定形时接受卫生部门的检查，一旦发现他们不可能治愈时，便使之无痛苦地死去.

1.8 历史上的色彩热

看了我以上的叙述，读者们一定会觉得，二维国里的生活实在是有些沉闷. 当然，我的意思不是说这里没有争斗、没有阴谋、没有骚乱或没有内讧等种种使历史显得热热闹闹的东西. 我也不否认，生命问题和数学问题在我们这里实现了奇特的混合；它们不断地激发出猜想和导致证实，使我们体验到你们在三维世界里简直无法理解的乐趣. 这里所说的生活单调呆板，是从美学和艺术的角度衡量的，而且实在是单调之极，呆板之至.

既然展现在我们每个人视野中的一切：风景也好，历史遗迹也好，肖像也好，鲜花也好，静物也好，统统都是些直线，只有明暗和模糊程度的变化，那还能不单调和呆板！

但情况也并不总是如此. 据传，在很久很久以前，我们也曾有过五六百年长的一段倏忽而逝的光辉岁月. 历史上曾有一位五边形 —— 其姓氏说法不

一 —— 在偶然情况下发现了若干种色彩的成分及着色的基本方法.据说,他先用彩色装饰了自己的房屋,然后是他厮养的奴仆,接着又给自己的父亲、儿子、孙子涂上了各种颜色,最后把自己也涂抹上了.由于着色一来不难,二者确实漂亮,人们无不为之所吸引.这位“夺目师”—— 这是权威人士考证出的姓氏 —— 所到之处,无不以自己的五彩斑斓的外观引起人们的注意和尊敬.此时,人们无需触摸,就能辨认出他,也不会把他的前后身搞错,他周围的人不用费心计算,便可判断出他在做什么动作.没有人会冲撞他,也不会给他让不成路.当没有颜色的四边形和五边形走在一群无知的等腰三角形之中时,总得要使劲发出声音来,以告知自己的存在,而夺目师却可以免去这项苦差事了.

这种着色热有如野火蔓延.不出一个星期,夺目师所在地区的所有三角形和四边形也都照此打扮了起来,只有少数特别保守的五边形仍不肯仿效.过了一两个月后,甚至连十二边形也受到了这一新潮的感染.未出一年,除贵族中的最高贵者外,各个阶层都形成了涂色的习惯;不消说,这种时尚迅速地发展到附近各地.经过两代人的时间后,除了妇女和神职人员外,二维国各地的人都是色彩绚丽的了.

至此,大自然本身开始对这种趋赶时髦的风尚表现出障碍了.对于愿意改革的人来说,长成多边形,这本身就是要求进行改革的一条根据.当时流行着这样一种风靡一时的说法:大自然造出许多边来,就是因为世界上存在着许多颜色的缘故.但是,这句话可显然不适于神职人员和妇女们.妇女只有一条边,而神职人员则一向自诩自己没有边 —— 他们认为自己是真正的圆,而不是生有无数多条无限小的高阶多边形.所以这两个阶层也可以认为上述改革的自然依据并不成立.因此,当其他所有的人都陶醉在用颜色装饰自己的身体而产生的魅力时,妇女和神职人员仍能“洁身自好”.

这场“着色革命”—— 也有人视之为无道德、无纪律、无政府、无科学的骚动 —— 从美学角度来看却标志着二维国艺术的光辉时期,但它未能进入成年阶段,甚至也未能迈入如花似锦的青年时代.生活在这一时期中的人真是快乐啊!色彩使他们的生活变得赏心悦目.即使是一个小小的聚会,也能搞得五彩缤纷.据说,当大家聚在一个礼堂或剧院里时,那富于变化的五光十色,竟使最出色的讲演人和演员的吸引力也大受影响哩.我又听说,最引人入胜的场面是阅兵式.

设想两万名等腰三角形齐刷刷地转身,把他们的黑色底边蓦地变换成橘红色或紫色的侧边;等边三角形则变换的是红、白、蓝三色;涂着紫红、佛青、橙黄和红棕四色的四边形炮兵在朱红色大炮的周围转来转去;而军医、测量官和副官等五边形和六边形们,也一会儿在这里露面,一会儿在那里现身 …… 难怪流传着这样一个故事,说有一位杰出的圆大人为他麾下的军队的艺术美所折服,





于是将元帅的权杖和王室的峨冠都抛到一边,大呼从今以后要与画笔为伍了呢!在那个时期,语言和词汇都有了很大的发展,这从一个侧面反映了当时美学的灿烂发展.

在“着色革命”时期,就连最普通的公民的最普通的言谈,也反映出较丰富的语汇和思想.我们最美的诗歌,以及保留在今日语言文学中的韵律感,都源自这一时期.

1.9 着色议案

但与此同时,各门非形象思维的学科却在迅速衰落.

着色时尚的出现,使视觉辨认术变得不再实用了.甚至在大学里,像几何学、静力学、动力学及相关学科,也都很快被视为多余之物而无人问津了.小学中讲授的较低级的触摸术也遭到了同样的命运.后来,等腰三角形阶层又宣称学校可不再需要什么活的小顶角标本,因此拒绝照常例送纳.过去政府的对策是一则以驯化,一则以减员,可现在,等腰三角形人口便激增起来,也更加桀骜不驯.

年复一年,士兵和手工业工人越来越强烈地宣称 —— 而且也越来越振振有词 —— 他们与最高阶层的多边形之间并没有什么重大差别,他们生来是彼此平等的;在有了着色术这个简单工具之后,他们也与多边形一样,能克服一切困难和解决一切生活问题,无论在静止条件下还是运动环境中都是如此.他们并不满足于视觉辨认术自然而然地被淘汰的前景,开始大胆要求用法律禁绝“垄断的贵族技艺”,并由此停止对视觉辨认、数学和触摸教学活动的所有资助.继而,他们又坚定地提出,既然色彩这个可称之为第二天性的存在已摧垮了划分贵族的需要,法律也应顺应形势,规定从今以后,社会各阶层的所有人等都是绝对平等的,大家的权利都应完全相同.

对此,高阶层的人举棋不定.这场变革运动的领袖们趁此又提出了进一步的要求,即所有的阶层应一律着色,神职人员和妇女也不例外.有人反对说,这两个阶层充其量只有一条边,无法着上不同的颜色.对此,鼓吹变革者们又提议让妇女把自己身体的前部——生有眼睛和嘴巴的部分——和后部用两种颜色区别开.为此,在一次特别全国代表大会上,有人提交了一份议案,建议妇女在身体的前半部分涂以红色,后半部分涂以绿色;神职人员也照此办理,其前半圆涂以红色,后半圆涂以绿色.

这是个极富计谋的议案.这不是等腰三角形提出的,而是出自一个不很规则的圆的谋划.由于社会上有人愚蠢地大发善心,使这个家伙童年时未被处死,这才导致后来这一议案的出台,给他的国家带来了祸殃,给无数同胞带来了死亡.

提出这个议案的动机之一，是把各个阶层的妇女争取到“着色革命”这一边来．革命的支持者们相信，由于妇女拟采用的两种颜色同神职人员相同，因此在一定的位置上看去，妇女会像是神职人员，这样便会得到尊敬的对待．对此，妇女们当然是不会不赞同的．

有些读者可能弄不清这一新议案为什么会造成神职人员和妇女的外观混同．我用几句话来讲讲清楚．

设想一名妇女照议案所说的装扮起来了，她的前半部（即生有眼睛和嘴巴的部分）是红的，后半部是绿的．从侧面上看去，可见到一段一半红一半绿的直线．

现在再设想一位神职人员，他的嘴生在点 M 的位置上（图 22.5），前半圆（AMB）是红色的，后半圆是绿色的，直径 AB 将红绿两色分了开来．如果你的眼睛与他的这条分色直径落在同一直线上来注视这位大人物，你也会看见一条直线（CBD），它的一半（CB）是红的，另一半（BD）是绿的；这条线也许比最高的妇女要短些，而且两端处也更模糊些，但它们二者在色彩上的一致会造成乍一看属于同一阶层的印象，从而不去留心其他细微之处的差别．考虑到这场着色革命已然造成了视觉辨认能力的退化，再加上妇女会用色彩模拟圆的明暗变化，读者们想必能够看出，《着色议案》中确实潜藏着使人们将神职人员同妇女混淆起来的危险．

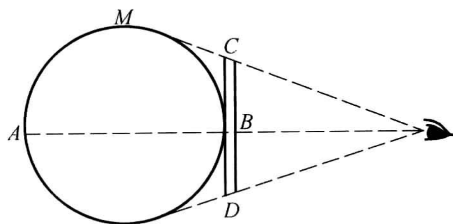


图 22.5

可以想象，这种前景对女人是何等地富有吸引力啊．她们高兴地看出，这种混淆会使自己捞到莫大好处：在家里，她们可以听到原来只为她们的丈夫和兄弟提供的政界和宗教界机密，甚至还可能假冒神职人员签署公文；在外面，她们的红绿两色完全同神职人员相同这一点，又能为她们伪装成圆招摇撞骗打开方便之门 —— 别人会把圆失落的东西恭恭敬敬地“奉还”给她们；她们的轻浮举止会让圆们背上黑锅 …… 至于国家大法因此会受到破坏，这些女人是不去管它的．即使是圆阶层的女士们，也都喜欢这个议案．

该议案的第二个目的是使圆们自身趋向逐渐腐败．在文化总的来说是江河日下的形势下，只有圆仍然保留着原有的纯洁和悟性，自孩提时候起，他们就很



少同色彩接触,因而保留了经过刻苦训练掌握的令人羡慕的视觉辨认技术,这样,在《着色议案》出台时,圆们不仅坚持了原有的立场,甚至还控制了这股时髦的风尚,加强了对其他阶层的管理.

我前面提到的工于心计的不规则圆是这一邪恶议案的真正炮制者.他要迫使高层人士接受这一滥着色的现状,从而一举降低等级制度的力量,取消高等阶层从小便能在家庭中接受视觉辨认术的机会,使色彩进入他们的家庭,使他们的智力萎缩.一旦圆也涂上颜色,他们的孩子就会在辨认父母亲时遇到问题——母亲经常被认作父亲,会使孩子对一切逻辑结论发生动摇.这样,神职阶层高智能的文化光辉将逐渐暗淡下去,整个贵族立法的毁灭和特权阶层的消亡便指日可待了.

1.10 镇压着色革命

《着色议案》引起的动荡达三年之久,到了后来,无政府状态似乎已注定要胜利了.于是多边形阶层便组织起一支武装,就连士兵都由他们本阶层的成员充当.然而,这支队伍被更强大的等腰三角形的武装歼灭了.在双方的战斗中,四边形和五边形保持中立.更糟的是,一些最有能力的圆也败在自己妻子的手下.在许多贵族家庭中,妻子们怀着政治目的,不停地纠缠男人们,要他们别再反对《着色议案》;有些人竟在劝说无效后,残暴地杀死了天真的孩子和无辜的丈夫,然后自己也一死了之.据记载,在连续三年的骚乱中,至少有二十三个圆的家庭毁于夫妻失和.

形势真是太危急了.神职人员似乎已到了除去屈从和灭亡之外别无选择的地步.就在这时,一件戏剧性的小事,突然使形势发生了根本的变化.(像这样的“小事”,政治家们决不应当忽略,他们应经常预料到,并有时还去促成之.因为这种事情特别能唤起民众的同情,其作用之大真到了匪夷所思的地步)

这件小事是由一个低等的等腰三角形引起的.他有一个很小的顶角,大约只有 4° 左右.他抢劫了一家商店,并用店里的颜色将自己涂抹成一个十二边形.(至于是他自己涂的,还是找别人干的,有关的说法不一)然后,他便去找一个他垂涎已久,但一直未能得手的贵族多边形的孤女求爱.由于一系列欺骗手段的偶然得手(因过程太长,这里不予赘述),也由于那位姑娘过于轻信,这个骗子成功了.婚后,可怜的姑娘发现受了骗,便结束了自己的生命.

这个悲剧在各地传开,猛烈地震撼了妇女们.对可怜的受害者的同情,以及对类似遭遇在她们自己或姐妹身上重演的担心,使她们对《着色议案》有了一个全新的看法.有不少人公开申明一改初衷,站到了反对者的立场上,其余的人只消稍加劝说,便也倒了过去.圆们抓住了这个有利时机,迅速召开了特别全国大会.大会上除了由在押犯担任警戒任务外,又请来许多妇女出席.

在这空前的集会上,圆阶层的领导人 —— 他的尊称为标准圆 —— 在起立发言时受到了十二万等腰三角形的各种嘘声.但他向与会者宣称,在大会结束后,圆阶层便将着手执行让步政策,遵从大多数人的意愿,并接受《着色议案》,从而使骚动立刻变成了喝彩.他邀请那位着色革命的领导者不规则圆来到会议大厅中央,请他以他那派人的名义,按受等级制度的终结.接下来,他用了几乎整整一天时间,发表了长篇演说.对这篇演说很难做出公平的评论,但可以肯定地说,它堪称修辞学上的杰作.他以公允的口气说,考虑到改革在即,应当对事情的全貌作一个最后的回顾:既看到《着色议案》的优点,也不忽略它的弊端.他一步步地指出了《着色议案》对商人、知识阶层和绅士的危害性.这时,等腰三角形们再次发出了不满的动静.于是,他又一次提醒这些人,尽管《着色议案》有许多缺点,但如果大多数人赞同,他还是愿意接受的.这一来,会场便重新安静了下来.除了那些等腰三角形以外,所有的人都被他的言辞感动了,他们或者表示中立,或者表示反对《着色议案》.

接下来,他又针对工人讲了一席话.他宣布工人的利益一定不会被忽视,又说如果他们有意接受该议案,至少应当对由此产生的后果有个全面的认识.他说,他们当中的许多人就要被擢升为规则三角形,其余的人也能指望自己的孩子争得父辈无望得到的地位.而现在,他们对荣誉的指望将随着《着色议案》的生效而落空,因为所有的区别将不复存在,规则形状与不规则形状将不可复识,进步将让位于倒退;再过几代人的时间,工人的地位将下降到士兵一级,甚至会到罪犯阶层,政治权力将落到最多数人,也就是罪犯阶层的手中,他们的人数比工人更多,而且将很快超过其他各阶层的人数之和.这一切都将是违背大自然法则的缘故.

一阵赞同的低语声穿过了手工业者的队伍.那个不规则圆觉得大事不妙,便试图上台讲话,但发现周围已经布了暗岗,于是只好噤声不语.这时,这位标准圆又说出几句诱惑力极强的话,把妇女们也给感动了.他是这样说的:如果《着色议案》得到通过,今后将不会有安全的婚姻,妇女也不会再有荣誉可言;欺骗、诡计和虚伪将浸透每个家庭,家庭的幸福将与宪法遭到同样的命运 —— 迅速地完蛋.说到这儿,他大声喊道:“在此之后,便是更迅速地来临的灭亡!”

这句话其实是一声行动暗号.会场内早已埋伏好了在押犯,听到这个信号便横过身来,一下子刺穿了不规则圆这个坏蛋.规则阶层的人群分了开来,为一群妇女让路,她们在圆的指点下走到前面,横过身来成为不可见的,然后百发百中地向士兵们戳去.工人们也像规则阶层一样让开道路,在押犯便一拥而上,以密集的队形守住了会场的各个出口.

这次战斗 —— 也许叫残杀更合适些 —— 很快便平息了.由于圆指挥得当,妇女们的刺杀发挥了很高的效能,而且许多妇女刺死了人,自己却连轻伤都未





受,还兀自在那里准备第二次厮杀呢!但这已经不需要了.那些等腰三角形的乌合之众自己来动手拆自己的台了.他们既失去了领袖,又受了看不见的攻击,后路又被在押犯切断,便立刻故态复萌,完全失去了头脑,嘴里“叛徒”、“奸细”地乱叫,将刚才的同盟者一下子当做了敌人,大家自相残杀起来.这便决定了他们的命运.只半个小时的时间,众多的暴乱者便无一幸存了.十四万彼此相残者的尸骸,无疑证明了秩序的胜利.

圆阶层不失时机地扩大战果.于是,十分之一的工人又被杀掉了;由等边形组成的民团也立即宣告成立;凡被怀疑为不规则形并多少有些根据的,一律无需社会部门的核准便可由军事法庭处决;军人和手工业工人的家庭受盘查达一年之久.在那个时期,全国每个城镇和村庄都系统地开展了清洗低阶层中过量人口的运动,他们一般是以往没有按规定送到小学和大学充当标本的罪犯,还有一些是违反了二维国宪法或其他法规的人.就这样,社会各阶层的平衡又得到了恢复.

不消说,颜色从此是再也不能使用和保存了,除了圆或者有资格的理科教师外,哪怕是在言谈中涉及任何同颜色有关的词语,都要受到严厉的处罚.据说只允许大学的最高级最深奥的某些课程中——这些课程我也无由参加——在阐发一些艰深的数学问题时有限制地使用颜色.至于是否确实,那我就不敢妄断了.

今天,在二维国的其他地方,一律都没有任何色彩.据悉,现在只有一个人懂得调色技术,他就是担任首脑的正规圆.这门技术只在他临死前才传给他的继承人.全国只有一个工厂生产颜料,那里的工人每年都被杀掉,然后再换上新工人,以免泄露秘密.连贵族阶层在回首遥远的过去因《着色议案》而引起的动乱时,还不免心有余悸呢!

1.11　神职人员

我该从有关二维国的概况,转到本书的主要内容,也就是空间的种种神秘之处了.这才是我的主旨,前面所讲的无非只是篇前言.

因此,我必须割爱许多内容.说实在的,要是讲这些内容,比如我们没有脚怎么能行走啦;没有手,又不能像你们那样打地基,也没有什么侧压力可以利用,却也能用各种建筑材料盖起房屋啦;我们这里各地带之间能保证北方地区不致截挡降到南方的水汽的降雨方式啦;我们的山脉、矿藏、树木、蔬菜、季节、收成都是怎样的情况啦;如何在看上去都是直线的条件下实现字母和文字的书写啦…… 凡此种种细节内容,读者们肯定是感兴趣的,但我都只好不讲了.我要提请读者注意,这样做的目的是为了节约大家的时间.

不过,在叙述最重要的内容之前,读者们一定还希望了解一下我们二维国

宪法的主要柱石,操纵芸芸众生的行为和命运的权威,受到普遍效忠乃至崇拜的等级 —— 圆阶层亦即神职人员阶层吧.

称他们为神职人员,并不是说他们同你们那里有这个称呼的人职能相同.对我们来说,神职人员是所有工业、艺术和科学的管理者,也是商业、贸易、军事、建筑、工程、教育、政治、立法、道德伦理和神学的总督导.他们自己从不做任何具体事务,但别人所做的一切该做的事情,无不出自他们的安排和命令.

他们被尊称为圆,这其实只是通俗化的叫法.受过良好教育的人都知道,所谓圆并非真正是圆,而只不过是有很多条边,而且每条边又都很短小的正多边形罢了.随着边数的增加,多边形就越来越近似于圆了.如果多边形的边数确实很多,比如三百或四百条,那么就是非常细心的人,也很难摸出他们的角来 —— 更确切地说,是即使让他来触摸,也很难摸出他们的角来,因为正如我在前面所说过的,上流社会是从不用触摸法进行辨认的,而触摸一个圆就尤其会被看做是大逆不道.最高阶层避免触摸的习惯更易使他们的真面目自幼便披上了一层神秘的面纱.一般多边形的周长为三英尺,有三百条边的多边形,每条边长便不会大于百分之一英尺,也就是稍大于十分之一英寸.而一个六百或七百条边的多边形,其边长只略大于你们三维世界里的一枚大头针的直径.当今的圆阶层领袖有一万条边.

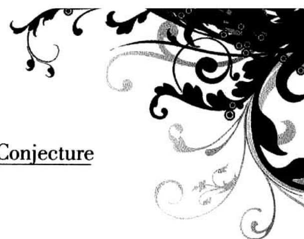
圆的后代在社会地位的提升上不受大自然的限制.自然法则只适用于普通的规则阶层,限定他们每一代只增加一条边.如其不然,成为圆就只是家世的积累这样一个数学问题了.也就是说,一个等边三角形的第四百九十七代孙,就一定是五百边形了.情况并非如此.大自然又确立了影响圆阶层无限增多的两条法规:第一,在社会等级上爬升得越高,边数增长得就越快;第二,生育能力同等级成反比.因而,四百边或五百边形很少能有子息,能有两个或更多儿子的这种家庭是绝对不存在的.另一方面,一个五百边形的儿子可能会有五百五十条边,甚至会有六百条边.

技术也来为高等阶层的进一步进化效力.医生们发现,可以通过手术使高阶层多边形的婴儿的柔弱的边折成两条更小的边,因此,将婴儿的每一条边精确地对折开,往往能使他的地位一下子提高几百代.当然,这种手术的风险很大,因此不常进行.

许多本来很有前途的孩子就在这种手术中夭折了,平均十个婴儿中仅能有一个幸存下来.然而,那些处在圆阶层外围的多边形父母都是如此地望子成龙,以至几乎无一例外地都把自己的幼子在不足月时送进医院来做这种倍边手术.

一年之内可见后果.在这个期限结束之前,多数孩子会被埋进墓碑林立的公墓.然而会有极个别的孩子经受住了考验,回到他狂喜的父母身边,而且不再是多边形,而是 —— 至少被称做是 —— 一个圆.为了指望得到这一幸福,有多





少多边形家庭付出了巨大的牺牲啊!

1.12 神职人员的教义

圆的教义可以一言以蔽之:悠悠万事,形状为大.无论是从政治、宗教还是道德方面,他们的教义都是为了个人和整个阶层在形状上的改进这一目的服务的.这表现在圆与妇女的关系上.

对整个社会来说,最不希望发生的事情是不规则形的出生.因此,祖上有不规则家世的妇女,决不是希望自己的后代能按升级规律在社会上步步青云直上的人心目中的理想伴侣.

一个男人的不规则是可以从外观上得到辨认的,但所有的妇女都是直线,看上去都是规则的,因此人们须想法搞清她们隐蔽的不规则性,也就是说,潜在于妇女身体内的、可能影响到其后代表现出不规则性的本性.有效的办法是由国家记录和保存好她们的家谱,不允许没有家世证明的妇女结婚.

因此,人们可能会觉得,一个圆,一个为他的家世骄傲,并希望有个能当上标准圆的后代的贵族,当然会比其他人更认真地选择家世无瑕的妻子.但事实却并非如此.一个人在社会上的地位越高,就越不看重女方是否是规则人家的女儿.对于一个极热衷于提高自己社会地位的等腰三角形来说,得到一个等边形的儿子的热望,会使他说什么也不肯找一个祖先中有过不规则记录的妻子.而四边形或五边形会鉴于自己的家庭地位能够稳定地上升,便不甚关心自己的第五百代会如何如何.六边形或十二边形对妻子的家世甚至会更加马马虎虎.而一个圆呢,三思之下选定的妻子,却可能有一个不规则的曾祖父 —— 她之所以被看中,可能只是因为她看起来比别人更光润些,或是生有迷人的柔美嗓音 —— 要知道,女人有动听的嗓子,是我们比你们更看重的一条审美标准.

不难料想,这样的不明智婚姻,结出的果实往往不是不规则的后代,就是边数减少的后代,甚至干脆就得不到果实.但迄今为止,这种不良结果仍未产生足够振聋发聩的力量.对于边数很多的多边形来说,减少几条边是不易觉察到的,何况还可以通过前面提到的那种倍边手术补救哩!

不过,如果不制止这种趋势,圆阶层人数的减少还会加快.那样的话,在他们当中不再能产生标准圆的时代就为期不远了,二维国的宪法也将随之完蛋.

还有个问题我觉得十分不妙,但又提不出什么补救办法.它也同我们与妇女的关系有关.大约在三百年前,当时的标准圆判定,妇女是富于感情的,但是缺少理性,所以不应将她们看做是有理性的生物,并由此规定不得使之接受智力教育.从此,她们就不会阅读了,甚至连计数丈夫和孩子角度的起码数学知识也不知道.她们的智能逐代下降.这一“女子无才便是德”的体制,或称之为放任自流的体制,今天仍然存在着.

我最担心的是，实行这一政策，已经给男性造成了损害.正是由于这样的政策，造成了男人的双重语言，甚至也可以说是双重思想.男人在女人面前大谈其“爱”、“责任”、“正确”、“错误”、“同情”、“希望”等非理性的情感类词语，其实，他们认为这些概念是并不存在的.将它们挂在嘴上，只是为了表现自己而已.到了男人之间，或者到了男人写书时，便又换成完全不同的语汇了，简直可以称之为“官话”——“爱”变成了“预期的利益”；“责任”变成了“必要”或“适合”；其他的词语也各有相应的转化.此外，当男人们置身于妇女群时，所用的语言会表现出对女士们的极大尊敬，这会使女人们相信，就是对于标准圆本人，男人们的敬意也未必会如此深切.可是一旦到了背后，她们在男人 —— 也许那些少男之辈除外 —— 的眼里简直就是摆设而已.

男人对妇女谈起宗教来，也同男人之间的谈话不可同日而语.这种语言上和思想上的双重训练，已成了强加在青年人头上的一种极为沉重的负担，特别是他们长到三岁后离开母亲身边，被教得不再使用原来已经学会的词语(当然，回到妈妈和保姆身边时还得再拾起来)，转而学习科学词汇和成语的时候更是如此.据我看，与三百年前相比，我们的祖先在对数学的准确掌握上反要比我们强.除了这一点之外，我们还不得不考虑这样的可能性：哪一天出了一个妇女，想法偷着读了书识了字，并把一些书中的道理讲给别的妇女听；或者有某个男孩子由于孝顺或由于不经心，对母亲道出了男人的这个秘密.就凭男性智力的衰退这一点，我也斗胆向最高当局呼吁，提请重新考虑有关女性教育的规定事宜.

2 一维国和三维国

2.1 我对一维国的访问

按我们的纪元算来，现在已经是 1999 年的倒数第二天了.这正是一个漫长假期的第一天.这一天里，我一个人做了不少有趣的几何习题借以消遣，一直干到很晚才去休息.夜里，我做了一个梦.

我看见我面前有极多的短线(我自然以为这是些妇女)，其间还夹有许多更短的，颇为润泽的小点.它们全都沿着一条直线来来回回地运动.据我看，它们的运动速度都相同.

他们的运动在一片嘁嘁喳喳声中进行，一旦运动停止，便一片安静.

我的位置正好靠近其中的一段最长的线.我想这是一位妇人，便走上前去同她搭话，可是没有人理睬我，再三发问也无济于事.我变得不耐烦了，便把嘴伸到她的前方位置上挡住了去路，又大声地重复了我的问题：“请问这位女士，





你在这条直线上不厌其烦地不停地走动,又发出这种奇怪的嘁嘁喳喳声,这是在干什么呢?”

“我可不是什么女士,”这条短线回答,“我是这里的国王.而你又是从哪里闯到我们一维国这里来的呢?”听了这出乎意料的答复,我连忙请这位国王陛下原谅我的惊驾行为,并恳求他向我这个外来客介绍一下他的国家.由于这位国王总以为他所熟悉的一切当然也是我都了解的,而我之所以总是问来问去,无非是为了开开玩笑而一个劲地装傻充愣.所以,我想要得到哪怕一点儿真正使我感兴趣的情况,都是十分困难的.不过我仍是百折不挠地问下去,终于了解到下述事实:

看起来,这是一位无知的可怜君主,他确信这条直线是他的国土;他在这里生活,而这里就是整个世界,就是全部空间.他只能在这条线上运动,只能看到这条线上的东西,此外他便一概不知、一律不懂.我第一次对他讲话时,他听到了我的声音,可在他听来同他以往的体验很不一样,于是便没有回答.他这样对我解释说:“我没有看到任何人,我觉得这些话像是从我自己的肚子里发出来似的.”直到我把嘴放到他的国土上的那一刻,他才算看到了我,也不再觉得有声音 —— 本来是来自外部,但他却觉得仿佛是来自他自己的肚子 —— 使他的身子产生振动感.对于我所在的二维地方,这个国王更是一点儿也闹不明白,在他的世界之外,也就是在他所在的直线之外,一切对他来说都是空白 —— 不,不能说是什么空白,因为空白也是空间,应当说是一切对他来说都是不存在的.

他的臣民 —— 那些短线是男人,亮点是妇女 —— 的运动和视野也同样被限制在这条线上(图22.6).不消说,他们的整个地平线只是一个点.每个人所能见到的一概是一个点:男人、女人、孩子、物体,在一个一维国民的眼里都是如此.他们只能靠声音来辨别对方的性别和年龄.此外,由于他们每个人都占了这条线上的一段地方,因此谁也不能给别人让路,这样一来,在一维国里决不会有

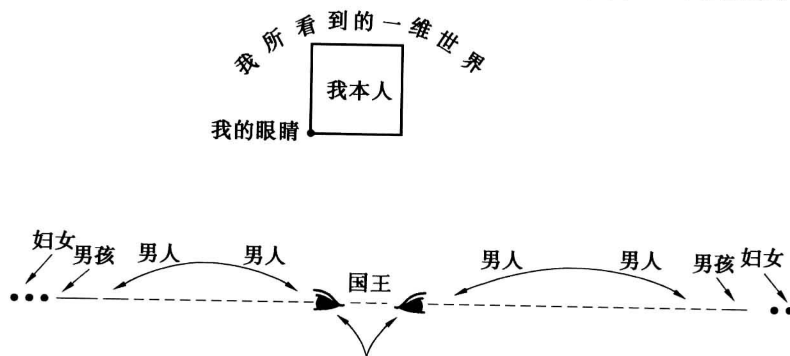


图22.6

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

人超越过别人.一旦是邻居,就永远得保持这种关系,这就像是我们那里的婚姻关系似的:只有死亡才能使邻居关系终止.

如此这种视野只是一个点,一切运动都离不开这条直线的生活,在我看来实在是说不出的乏味.但我惊奇地注意到,国王却是蛮愉快的.我想,在这种糟糕的环境里,怕是无法建立私人交往,也不能享受夫妻生活的吧?要不要向国王提出这个微妙的问题呢?踌躇再三之后,我还是启齿问了他家庭成员的健康状况.对此他答道:"我的妻子们和孩子们都很好、很幸福."

我觉得很奇怪.在进入一维国、靠近国王陛下之前,我已经注意到,国王的左右只是些成年人.于是,我又发问道:"请原谅,我真想象不出,王后陛下和王子、公主殿下能在什么时候以什么方式看见和靠近您.喏,你们之间至少隔着六个人,而你既不能从他们的身子里钻过去,也不能贴着他们滑过去.难道在一维国里,人们的婚配和生育无须凑到一起就能进行吗?"

国王回答说:"你这个问题有多么荒唐!如果真的像你想的那样,世界上很快就会一个人都不剩了呢!不,不,不,两颗心的结合不必以相邻为先决条件,生儿育女这种大事,更不能让偶然的邻里关系左右.你可也太无知了.不过,既然你说了这样无知的话,我也就真的把你当做一维国里最天真的人指教一下吧.记着,婚姻是靠声音、靠听觉缔结的."

"你当然知道,男人除生有两只眼睛外,又都生有两张嘴巴、两副嗓子.它们长在身体两端的位置上,一头是低音,一头是高音.本来我用不着提到这件事,但在刚才的谈话中,我为什么听不到你的高音呢?"我告诉他,我只有一副嗓子,也不知道他的竟是两副.国王又说:"那就越发证实了我刚才的看法,即你不是个男人,而是个生着粗嗓门,耳朵也完全没有经过训练的畸形女人.不过,我还是接着说好啦."

"大自然有它自己的法则,那就是每个男人要娶两个妻子……"

"为什么?"我忙问道.

"你装蒜也装得太过分了吧!"他叫道,"没有四者为一的结合,即一个男人的男高音和男低音同一个高音女人和一个低音女人的结合,怎么能实现完全和谐的整体呢?"

我又说:"也许,有的男人宁愿要一个妻子,或者希望娶三个老婆呢?"

"这是不可能的,"他说,"就像二加一等于五或者像一个人能看见一条直线一样地不可能."我正想再次打断他的话,他又继续说了下去:"在一条自然法则的驱遣下,我们每个星期会有一次做出节奏比平时更激烈的往复运动,时间会持续一段从一数到一百零一那么长的功夫.大家一边唱一边动,到了该数第五十一下时,人们就停止了一切活动,每个人都发出自己最华丽、最丰满、最甜蜜的声响.婚姻正是在这一决定性的时刻缔结成的:一个男低音配上一个女





高音,一个男高音配上一个女低音.爱人之间常常相距几万里之遥,但照样能听到并应答冥冥中注定会使自己中意的情侣的声音,距离上的障碍其实是微不足道的.每三个相爱的人结合为一体,其圆满的结果便是一维国里会增添许多男孩和女孩,每一家一次生出两个女孩和一个男孩."

"什么?总是三个一组吗?"我问,"这就是说,两个妻子中的一个必须生一对双胞胎吗?"

"是的,你这个低嗓门的畸形女人!"国王答道,"不然的话,又怎么能保持一男两女的比例平衡呢?你难道不懂得最基本的自然法则吗?"他生气地不说话了.又过了好一会儿,我才又引得他张嘴讲话.

"当然,你不会认为,在这个婚姻大合唱中,每个单身汉都能靠第一次引吭高歌缔结良缘.相反,我们之中的大多数人要唱上好多次才成.只有少数爱情的幸运儿能立即相互一听倾心,从而一下子找到上帝给他们安排好的配偶,完美地结合成和谐的一体.大多数人的求婚期是相当长的,有的求婚者的声音也许只能与一个女人的声音相融洽,也许一开始时同谁都不协调,还可能是一个男人的两个未婚妻发出的女高音和女低音之间不很协调.在这些情况下,靠着大自然为我们提供的每周一次的大合唱,相爱的三者在声音上便逐渐地向和谐靠拢了.每试一次高歌,每发现一个不和谐的和声,都会促使有缺陷的人努力向完善方向发展.经过这样多次练习和修正,最后才能取得成功.总有一天,当奇妙婚姻大合唱再次在整个一维国里响起时,遥遥相望的三者间便忽然发现他们的声音是极其和谐的,于是,他们便在销魂的和声之中结合了.大自然便会高兴地看到又一桩好姻缘和三个小宝宝的出世."

2.2　向国王解释二维国的一番徒劳

我觉得该让这位国王从自己那沾沾自喜的认识中清醒一下了.于是,我决定费点工夫,让他多少看到一些真理之光,也就是说,让他了解一些有关二维国的情况,我是这样开始的:

"陛下是如何确认你的臣民的形状和所在位置的呢?在进入你的国土之前,我看到你们有些人是线,另一些人是点;有些点较长……"

"你都胡扯些什么呀!"国王打断了我,"那一定是你的幻觉.每个人都知道,靠视觉来分辨线和点是绝对不可能的.但我们可以凭听觉推断出这一点来,而且这一方法还可用来精确地确定一个人的形状.就拿我来说吧.我是一维国里最长的线,占有六英寸的空间……"

"长度."我贸然纠正道.

"傻瓜,空间也就是长度呀!你要再打岔,我就不跟你说了."

于是我向他道歉,他这才接着以轻蔑的口气说:

“和你讨论真是费劲.让找来告诉你,我是如何用我的两副嗓子向我的妻子们表示我的形状的吧.她们此刻距我有六千英里零七十码两英尺八英寸之遥;一个在南,另一个在北.听,我向他们呼叫了.”

他发出一阵嘁喳声,然后继续对我说:

“我的妻子们此刻已经听到了我用第一副嗓子发出的声音,而用另一副嗓子发出的声音,她们紧接着也会听到.她们会觉察到,在这两个时间间隔里,声音能传播6.457英寸远,因此可推出从我的一张嘴巴到另一张嘴巴之间的距离就是6.457英寸,于是便知道了我的形状就是6.457英寸长.当然啦,我的妻子们是用不着听一次就计算一次的.在我们结婚之前,她们已经计算过了,这是一劳永逸的.用同样的方法,我也能靠对声音的分析得知我国任何一位男性公民的形状.”

“可是,”我说,“如果一个男人只用他的一副嗓子发音,并且装出女人的声音来,不然就用一副真嗓一副假嗓出声,让别人听不出来它们是同一个人发出的呢?这种欺骗不是会造成很大的麻烦吗?其实,你只要命令相邻的百姓间相互触摸一下,不就能检查出这种欺骗行为吗?”当然,其实我也知道,触摸在这里是不解决问题的,这当然是个愚蠢的问题.我这样问的目的,就是想惹恼国王,结果真的成功了.

“什么!”他大惊小怪地叫道,“你这是什么意思?”

“触摸呀,就是接触一下,也就是彼此挨一挨身体.”我解释道.

“如果你的意思是接触,”国王说,“也就是让两个人靠到一起而不留有任何空间,那么陌生人,我告诉你,在我的统治下,干这种勾当的人是要处死的.原因很明显,这种接触会使妇女纤弱的身躯破碎,因此须由国家加以保护.由于视觉不能区别男女,所以法律规定,无论男人或女人,均不得接近到两者之间没有间隔的程度.”

“而且,说真的,即使进行你所说的‘接触’,也就是违法地进行逆悖天然的接近,又能达到什么目的呢?凡用你所意指的这种野蛮粗鲁的过程所能实现的目的,都能凭听觉更迅速、更准确地达到.至于你所说的上当受骗的危险,也是根本不存在的.一个人的声音是天生的,不能随意改变.好了,我们回过头来接着说吧.即便我能穿透固体,那么,我的臣民有上亿个,我就得一个接一个地从它们的体内穿过去,好靠触觉了解他们的大小和距离.这种方法既笨拙又不精确,而且又很浪费时间和精力.可现在,我只要听上一阵子,问题就都解决了.听就是我的人口统计,无论是地区统计,人头统计,还是智力统计,我都能一个不漏.这一切都靠听,只靠听!”

说着,他又收住了话头听起来,好像都听入了神,可是我却只觉得听到一片细小的嘈杂声,并不比一大群超小型蚱蜢发出的动静好听多少.





我回答说:“真的,你的听觉对于你是很有好处的,它能弥补你的许多不足.但请允许我指出,你在一维国里的生活肯定是可悲地单调乏味,除了一个点,你们什么也看不见,甚至连一条线也无法看到 —— 不,甚至都不知道什么是线!你们甚至被剥夺了看到线的权利,而上帝把它赐给了我们二维国人!要是只能看到区区一个点,那真不如干脆不长眼睛的好!确实,我并没有你们那种精细的听觉能力,所以,叫你听得如醉如痴的一维国音乐,在我听来无非是一片嘁嘁喳喳.但我甚少可以用视觉来分开一条线和一个点.我现在就再向你证明一下.就在我进入你的王国之前,我看见你先从左到右,后来又从右到左地跳着舞,在你左边有七个男人和一个女人离你很近,在你右边则是八个男人和两个女人.对不对?”

国王说:“就数量和性别来说是对的,虽然我不知道你所说的‘左’和‘右’是什么意思.可我肯定地说,你不是看见他们的.你怎么能看见线呢?这不等于说你能看见一个男人的身子吗?你一定是听出来的,只是梦见自己看到了这些罢了.现在,让我来问问你,你的‘左’和‘右’是怎么回事?你是说‘南’和‘北’吧?”

“不,”我回答,“除了你们向南和向北的运动之外,还存在另外一种运动,我就叫它为从左到右的运动.”

国王:“如果你愿意,是不是让我具体看一下这种左右运动?”

我:“不,我做不到,除非你能从你的直线上走出来.”

国王:“走出我的直线?你的意思是要我走出世界,走出空间?”

我:“是的,走出你的世界,走出你的空间;而你的空间并非真正的空间,真正的空间是个平面,而你的所谓空间只是一条线.”

国王:“如果你不能给我表演这种左右的运动,那就用语言描述一下也成啊.”

我:“如果你不能确定自己的左右,我恐怕也没有什么办法说清楚.其实,这么一种简单的区别,你肯定是不会不知道.”

国王:“可我就是压根儿也不知道呀!”

我:“天啊!我怎么才能使你明白呢?当你在直线上运动时,是否有时会想到,你可以向别的方向运动呢?用你的眼睛向你的身体所直对的方向看一下吧.要么换个方式说一下,除了向你的两端方向运动之外,难道你从来不曾想要改变一下你的运动方向吗?比如说,难道你不想向你侧面的方向动一下吗?”

国王:“从来没有.再说,你这又是什么意思?难道一个人的身子能对着什么方向吗?一个人怎么能朝着他身子的方向运动呢?”

我:“好啦.我用语言说不清楚,还是用行动来试试吧.我要慢慢地离开你这个一维国,以此向你表演一下我沿着要向你说明的方向的运动.”

我一边说着,一边把身体渐渐移出了一维国.在我的身体仍有一部分留在他的国土上的整个期间.这位国王一直在叫喊着:"我能看见你…… 我还能看见你…… 你没有动啊!"可是,当我最终把身体全部移出了他那条直线之外时(图 22.7),他刺耳地尖叫起来:"瞧!她不见了,她死了."我回答说:"我并没有死,只是离开了一维国,也就是说,离开了你称之为空间的直线.此刻,在这真正的空间里,我能看见事物的本来面目.我看见你是一条直线,看到了你的身子,也能看见在你北边和南边的那些男人和女人.我就来点一点他们的数目,说出他们的排列顺序,以及他们的大小和间距."

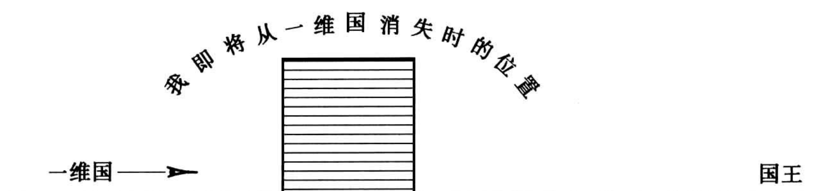


图 22.7

我向他表演了这一切,然后以胜利者的口气问道:"现在你总该相信了吧?"一边说着,我一边再次进入一维国,还停在我刚才所在的位置上.

可这位国王却对我说:"如果你是一个有理智的人,哪怕只有一点点理智,你就该先听后推理.你要我相信,除了我感觉到的这条线之外,还有别的线存在;除了我平时意识到的这种运动之外,还有别的运动可能.那么请问,你能不能用语言来描述一下你所说的别的线呢?其实刚才你根本就没有动,只是搞了出大变活人的戏法.你并没有对你讲的新世界给出任何实在的描述,却只是告诉给我我周围的随员的数量和大小.这可是我们都城中连孩子都知道的事实咧.还有什么能比这一套更胆大妄为和荒谬绝伦的呢?承认你的愚蠢吧!否则就从我的国土上滚出去!"

由于他的刚愎自用,更由于他竟声称不知道我的性别,我于是发起怒来,嘴里便没有遮拦了:"你这痴迷不悟的东西!你自以为生活得挺美,其实你是个顶不怎么样的低能儿.你夸耀你能看得见,但只能看见一个点!你自诩能推断出一条直线的存在,而我却能真的看见一大堆,还能推断角的存在,以及三角形、四边形、五边形、六边形甚至圆的存在.我不想再浪费口舌了,只想再对你说一句:我是完善的,而你是不完善的;你是一条线,而我是由无数条线组成的,在我的国家称做四边形.我远比你优越得多,但同我那里的大贵族相比,却只是一个微不足道的人.我来自二维国,本打算给你来点儿启蒙教育……"

听到这里,国王威胁地喊着向我走来,仿佛要斜向将我刺穿似的.与此同时,他的无数臣民也发出一片作战的鼓噪.这片呐喊声越来越强,简直使我觉得



有如十万名等腰三角形士兵的怒吼和五边形炮兵用一千门大炮发出的轰鸣.我迷茫地待在那里,既不能说,也不能动,只觉得小命快要玩儿完了.声音越来越响,国王离我越来越近了…… 于是我便醒了过来,耳边正响着二维国里催我前去吃早饭的铃声.

2.3 来自三维国的陌生人

从梦中醒来,我又回到了现实.

按我们的纪元算,已是1999年的最后一天了.滴滴答答的雨声早就宣告了夜晚的降临.我坐[①]在妻子身边,回想着一幕幕往事,也展望着未来——未来的一年,未来的一百年,未来的一千年.

我的四个儿子和两个失去父母的孙子都已回到他们各自的房间去了,只有我的妻子留下来,和我一起等待着第二个一千年的消逝,迎接第三个一千年的到来.

我仍然反复思考着刚才我的小孙子偶然说出来的话.他是一个非常聪明的孩子,是个非常完美的六边形.我和他的叔叔们一起教他视觉辨认术.我们绕自己的中心点时快时慢地转动,让这个孩子说出我们的位置.他的回答总是令人满意.为了鼓励他,我又教给他一些几何学上有用的数学知识.

我拿来九个四边形,每个边长均为一英寸,我把它们放在一起,便组成了一个边长为三英寸的大四边形.由此我向我的小孙子证明——因为我们无法看到它的内部——将每个边上的英寸数平方一下,就会知道它的面积有多少平方英寸.我告诉他:"这样,我们就能知道,这个边长为三英寸的四边形面积的平方英寸数是九,也就是3^2."

我的小孙子思索了一会儿,便对我说道:"你教我算过一个数的三次方,我想,三的三次方3^3一定是有几何意义的,但是有什么意义呢?"

我回答道:"什么意义也没有,至少没有几何意义.几何学里只存在两个维."然后,我又对他演示了一个点怎样通过沿一条线移动三英寸而生成一条三英寸的直线,在数学中以3代表之;然后,这条三英寸的线又平移三英寸,便得到了每边为三英寸的四边形,以3^2表示之.

对此,我孙子又拾起了他原来的想法.他一下子打断我,兴奋地喊道:"这就对了!如果一个点移动三英寸便生成一条三英寸长的直线,并可以以3表示;一条三英寸的直线平移后生成边长为三英寸的四边形,并可以以3^2表示,那么,这个边长为三英寸的四边形若再以某种方式平移一下——怎么平移,我可就

① 当然,我说的"坐"其实并没有你们在三维国里的表示某种姿势的含义.我们没有腿,因此既不能"坐",也不能"站",就像你们那里的比目鱼也不能坐和站一样.但我们也使用"躺"、"坐"、"站"这几个词语,它们是用来表示愿意程度和精神状态的.它们代表的愿意程度依次减弱.不过,对于这个问题,以及与之类似的千百个问题,因为时间关系,我就不再赘述了.

说不出来了 —— 就一定会生成一个每边长三英寸的什么东西 —— 我也说不出是什么东西 —— 而且一定可以以 3^3 表示 ……”

“睡觉去吧!”对他的插嘴,我有点生气了,“如果你少来点儿胡说八道,就会多记得些道理.”

于是,我的孙子只得快快离去.我坐在妻子身边,竭力回顾 1999 年,展望 2000 年.可这个聪明伶俐的小家伙的想入非非却老在我的脑子里打转,怎么也摆脱不了.半小时一颠倒的计时沙漏里还有不多的沙子了.我从沉思中惊觉过来,在第二个一千年的最后时刻里最后一次把沙漏转向北方,同时生气地大声说道:“这孩子真是个小傻瓜!”

突然,我觉得房间里好像又有了一个人.顿时,一股冷气袭遍我的全身,使我感到毛骨悚然.

“他可不是傻子!”我的妻子喊道,“这样说你孙子的坏话,不是违背戒条的吗?”

但我并没有注意她的话,只顾向四周打量,却没发现任何人,但我还是觉得真的有个人,似乎我又听到什么声音.我一下惊跳起来.

“你是怎么啦?”妻子问道,“没有穿堂风啊!你要找谁?这儿谁都没有啊!”

的确什么都没有.我又回到老地方,再次大声说:“我说这孩子是个傻瓜,3^3 没有几何意义.”话音刚落,我便清清楚楚地听见有个声音回答:“这孩子可不傻,3^3 有明显的几何意义.”

我和妻子都听见了这句话 —— 当然,她并不明白是什么意思.我们都跳起来,向有声音的方向张望过去.我们看见面前出现了一个图形.对此,我们是多么恐怖啊!乍一看这是个女人,正侧向对着我们,但再瞅一眼,就能明显地看出,这个人有极迅速地变得模糊起来的两侧,因此又不像是位女性.我又猜想这是个圆,但又好像是个能改变大小的圆,而根据我的经验,无论圆还是其他任何规则图形,都不可能以这种方式变化.

我的妻子可不如我有经验,也不像我那样会冷静地注意观察.由于女性的急躁和没来由的嫉妒,她立即断定这是一个女人从哪个小缝隙里钻了进来.她生气地喊:“这个人是怎么进来的?亲爱的,你可是答应过我,在这所新房子里不开通气孔啊!”

“是没开通气孔,”我说,“可是,你为什么就认为这个陌生客是个女人呢?依我用视觉辨认术判断,我看见的是 ……”

“哟,我可没工夫听你那套视觉辨认的玩意儿,”她回答我,“要知道,谚语里说‘耳听为虚,触摸为实’、‘看时一个圆,摸时是条线’啊.”(这两条谚语是二维国女性们常挂在嘴边的)

“好啦好啦,”我不想惹恼她,便说道,“如果真是这样,那咱们便同这位客





人彼此认识一下吧.”我妻子便以最优雅的仪态走向陌生人说:“这位女士,请允许我请求您摸一摸我,也让我……”话未说完,她忽然退了回来:“哟,这不是个女人,而且也没有角度,压根儿也没有!我刚才是不是无礼地摸了一位标准圆呢?”

“从某种意义上说,我的确是个圆,”这个声音回答,“而且比二维国的任何一个圆都更完美.更确切地说,在我身上有许多个圆.”他又更温和地加了一句:“亲爱的夫人,我有件事要同你丈夫谈,但不便当着你的面讲.劳驾您离开一会儿……”我的妻子不肯让这位尊贵的客人纡尊降贵地请求,忙说自己早该回屋休息了,并又就刚才的失礼道歉不已,最后总算真的走了.

我瞥了一眼沙漏,最后一点沙子已经漏下去了,2000年——第三个一千年的第一年——开始了.

2.4　陌生人向我揭示三维国奥秘的一番徒劳

我妻子的告别声刚刚消失,我就开始靠近这位陌生人,打算请他“坐”下来,并借此更近地看看他.一见他的外表,我就目瞪口呆、痴若木鸡.他竟连一点点带角度的征候也没有,但他的大小和光亮程度又每时每刻都在变化着.据我所知,这在任何图形来说几乎都是不可能的.我突然想到,这可能是一个窃贼或强盗,是装成圆的某个邪恶的不规则等腰三角形,不知用什么方法钻进了我的屋子,正打算用他的尖角刺穿我呢!

但我是在房间里,而且目前正值干燥季节,所以没有雾气,这使我很难相信视觉辨认的结果,特别是我现在就站在离他极近的地方.我恐怖极了,竟唐突地向他冲去,说道:“先生,你必须允许我……”一边说,我一边触摸起他来.

我的妻子是对的,他没有角的迹象,也没有一点有棱角或不均匀的地方,在我的生活中,从没有见到过比他更完美的圆了.我从他面前开始,围着他转一圈.当我绕着他走动时,他一动也不动.我看出他的确是个圆,一个极好的圆,这是毫无疑问的.

后来便是一番谈话.我在这里尽量将原话复述出来,只是把其中没完没了地道歉的话砍掉了一些——我竟斗胆摸了一个圆,事后想起来真是羞愧难当啊.谈话是以来客对我长时间的触摸表示不耐烦开始的.

陌生人:“现在你该摸够了吧?可你还没有自我介绍呢!”

我:“最杰出的阁下,请原谅我的唐突,这不是因为我不懂得文明社会的规矩,而只是惊奇和紧张使然.您的造访实在是太出乎我的意料了.阁下,我恳求您不要把我的失礼行为告诉别人,特别是我的妻子.在和您交谈之前,请您满足一下我的好奇心,让我知道您是从哪里光临的,好吗?”

陌生人:“从空间.我从空间来,先生.不从空间,又能从哪里来呢?”

我:“对不起,阁下.不过就在此刻,您和鄙人不也是在空间里吗?”

陌生人:“非也,非也.你认为什么是空间呢?你能下个定义吗?”

我:“空间就是无限延伸的长和宽,阁下.”

陌生人:“说真的,看来你连什么是空间都不知道呢.你以为空间只是二维的吧?我再告诉你第三个维 —— 除了长和宽以外,还有一个高.”

我:“阁下可真爱开玩笑.我们也把长说成高,把宽说成厚,因此是两个维的四种说法.”

陌生人:“我的意思并不是叫法问题,而是说存在着三个维.”

我:“请您说明一下,这个我所不知道的第三个维在什么方向上?”

陌生人:“就从我来的那个方向上,也就是沿着上和下的方向.”

我:“您的意思是说沿着南和北吧?”

陌生人:“根本不是这个意思.我所指的方向你是看不见的,因为在你的身子边上没有生着眼睛.”

我:“请原谅,阁下.您看一下就会相信,我是生有眼睛的.在我的两条边的相交处有一个发亮的小点,那就是我的眼睛.”

陌生人:“我知道.但为了向空间看,你应当还生有别的眼睛.不是在你的周边上,而是在你身体的侧面.可能你会叫它是你的体内,可在我们三维国,则称之为侧面.”

我:“让我体内长一只眼睛?长在我的肚子里?阁下是在开玩笑吧!”

陌生人:“我可没有开玩笑的闲心.我跟你说过,我是从空间来的,可你不懂得什么是空间,所以我又对你说,我来自一个有三个维的国度,我刚才就从那里俯瞰了你们自称为‘真正的空间’的这个平面.从刚才那个有利的位置,我可以进一步描述你们称之为‘立体’(在你们这里,凡是四面都有边界的,你们都叫它是立体)的东西:你们的住房、教堂、家具、保险箱,甚至你们的内脏,全都暴露在我的视野里.”

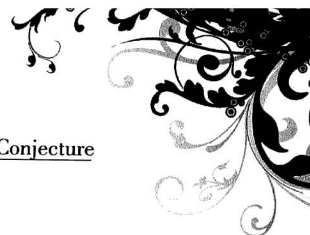
我:“嘴上这样说说倒也不难,阁下.”

陌生人:“你的意思是说证明起来可就难了不是?可我这就来证明一下.当我往这儿下降时,我看见了你的四个儿子,他们是五边形,都在自己的房间里;还有你的两个六边形的孙子.我看见你最小的孙子和你在一起呆了一会儿,然后回到了自己的房间,这间房子里只留下了你们夫妻二人.我还看见你的三个等腰三角形的仆人在厨房里准备晚饭,另有一个小听差在帮忙洗碗碟.然后我便到这里来了.现在,你认为我是怎么来的呢?”

我:“我想,您说不定是从天花板的什么地方钻过来的.”

陌生人:“不对,你很清楚,你的屋顶最近刚刚修缮过,甚至连个能让女人钻过的缝隙都没有.告诉你,我是从空间来的.我已经对你讲了有关你的孩子和家





庭的情况,你还不相信我是从空间来的吗?”

我:“阁下一定很清楚,像您这样一位能拥有广泛情报来源的贵人,要想从我周围了解到一些鄙人家里的私事,不是易如反掌吗?”

陌生人:(自言自语)“我该怎么办?对了,我又想起了一招.—— 当你看一段直线,比如看你的妻子时,你认为她是几维的呢?”

我:“听阁下同我谈话的口气,就好像我是个粗人,对数学一无所知似的,所以觉得我会认为女人真的是一段直线,只具有一个维.不,不,不,阁下.我们四边形多少是喝过些墨水的,因此能像您一样,知道妇女虽然一般被称做直线,其实科学地说是一个极细长的平行四边形,因此有两个维,同其他人一样有长有宽.”

陌生人:“正因为有三个维,你才能看见这条所谓的线呢!”

我:“阁下,我刚刚说过,妇女是有宽度和长度的.我们能看见她的长,推断出她的宽,宽度虽然极小,但也能测量出来.”

陌生人:“你没有明白我的意思.当你看见一位妇女时,你应该 —— 除了推测她有宽度之外 —— 看见她的长,也应当看见她的我们称之为“高”的东西的存在,虽然最后这一个维在你的国家里是无穷小的.如果一条直线只有长而没有高,它就不再占有空间而变为不可见的了.你真的没有认识到这一点吗?”

我:“我承认我一点儿也不明白您的意思.我们在二维国看一条线时,是凭借长度和亮度看见的;如果光亮消失了,就说明这条线已不复存在了,也就像是你说的不再占有空间了.据我猜想,你所说的另一个维,就是这个我们称之为‘亮度’的东西吧,只是你们称之为‘高’就是了,对吗?”

陌生人:“不.我所说的高,就像你们的长度一样,也是一个维,只是在你们这里,高度一概是极小的,因而不易觉察出来罢了.”

我:“阁下,对您所说的是很容易加以验证的.您说我具有你们称之为‘高’的第三维.那好吧,维是有方向和大小的.那就请您测量一下我的‘高’,或者指一下我的‘高’是沿什么方向延伸的吧.如果您能叫我知道,我就会信服您,否则就只好对您所说的敬而远之了.”

陌生人:(自言自语)“既测量不出来,也无法对他指出?怎样能使他信服呢?有了,先对他说几个事实,然后再实际演示一下让他看看,这肯定就能行了.—— 请再听我说,先生.

“你生活在一个平面上.设想你称之为二维国的世界,就是人们称之为流体这个东西的广大的表面,你和你的国民们就在这个面上,或者说在这个面里活动,但是不能上升或下降.

“我不是平面图形,而是一个立体.你称我为圆,可实际上我不是圆,而是叠在一起的无数多个直径从零至十三英寸的种种大小不等的圆.当我切入你的平

面时,就会在这个面上截成一个图形,这就是一个圆.瞧,我现在就是这样做的.因为对于一个球来说——球就是我在我那个国家里的名称——如果能向二维国民表示出自己的形状的话,也只能表现为一个圆.

“你记不记得你昨夜的一维国之行?—— 我是什么都能看见的,因此我已经看到了你脑子里存留的有关幻象.喂,你记不记得,在你进入一维国时,是如何不得不在国王面前显现为一条线而不是四边形的?那是因为一维国只有一个维,因而不足以表示出你的全貌的缘故.现在的情况也完全类似.你们二维国也同样没有足够的地方来让我表现出自己的全貌;我是个三维形体,因此在这里只能显现我的一个断面,这也就是你称之为圆的图形.

“你的眼神有点暗淡,这表明你仍不肯相信.但现在我就要用确凿的证据表明,我告诉你的都是真理.既然你没有能力把你的眼睛抬出你的平面之外,你就只能看到我的一个断面,也就是说,每次只能看见一个圆.可当我在空间上升时,你至少会看见我在平面上的断面是逐渐变小的.现在你来看,我升起来了 …… 你看见的圆正变得越来越小,直到缩成一个点,并且最后消失.”

我看不见他的上升,只是看到这个圆一点点变小,最后便消失了.我眨了好几次眼睛,为的是弄明白是不是在做梦.这不是梦.从空中不知何处传来了缥缈的声音 —— 它好像就贴在我的心口:“我离开了吧!好,现在我再渐渐回到二维国来,你会看见我的断面越来越大.”

在三维国的每个读者都很明白,我这位神秘的客人所讲的话明摆着都是真实的.但我虽然说是二维国里一个精通数学的人,可这对于我却不是那么简单明白的事情.三维国里哪怕随便一个孩子,看了上面的那张很大略的图解,都会明白这是一个球在三维空间里上升时,对我或任一个二维国民表现为一个圆,而且开始时最大,然后就变小,最后变得极小,近似于一个点.而我却虽然看见了眼前的事实而依然感到莫名其妙.我所能理解的,只不过是这个圆使他自己变小和消失,然后又重现并迅速变大.

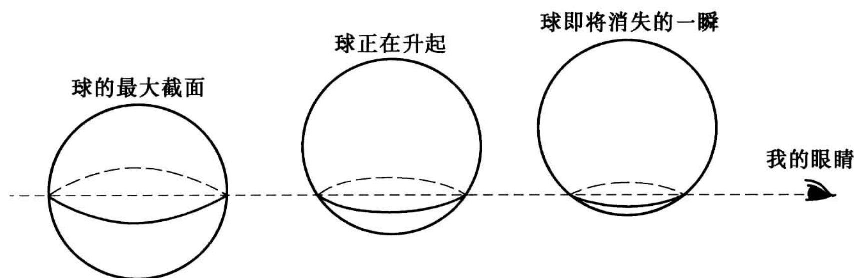


图 22.8

当他又恢复到原来的大小时,大声地叹了一口气,因为他从我的沉默中觉察到我仍完全不理解.确实,现在我倾向于相信他全然不是个圆,而可能是个高



明的魔术师,要不然就是真像有些老婆婆们所说的那样,世上毕竟还有巫师方士之类的人.

又沉默了许久之后,他又自言自语了:“如果我不靠行动证明的话,至少还能有一个办法,那就是试一试类推法.”随后,他又同我讲起话来.

球:“数学家先生,请告诉我,如果一个点向北移动,留下一条发亮的轨迹,你称这条轨迹为什么?”

我:“一条直线.”

球:“一段直线有几个端点?”

我:“两个.”

球:“现在设想这条线沿东西方向移动,它的每一个点所产生的轨迹都形成一条直线,这样形成的图形叫什么?我们假设这段线平移的距离与其本身等长.说说看这叫什么?”

我:“一个四边形.”

球:“一个四边形有几条边?几个角?”

我:“四条边、四个角.”

球:“再开动一下脑筋,想象一下在二维国里有一个四边形一致地向上移动它自身.”

我:“向哪里?向北吗?”

球:“不,不是向北,是向上,完全脱离二维国.

“如果这个四边形向北移,它在南边的点必定会通过北边的点原先所占据的位置.这个不是我的意思.

“我的意思是你身上 —— 因为你就是个四边形,所以我就以你为实例来说明 —— 的每一个点,也就是你认为位于你身体内部的每一个点都向上通过空间,使你身上没有一个通过其他一个点原先所占据的位置,但每一个点本身都描绘出等于自己的一条直线.这正是地地道道的类推法.你一定明白了吧?”

现在我真感到一种强烈的冲动,想不顾一切地冲向这位来访者,把他抛出去,把他赶出二维国,把他弄到随便什么地方去,以使我能摆脱他.但我还是极力克制住了自己的厌烦情绪回答道:“你乐于用‘向上’这个字眼来表示的这种运动使我形成的这个图形会有什么性质呢?我想,用我们二维国的语言是可能描述出来的吧?”

球:“当然,这一切都是简单明了的,而且完全可以类推出来.顺便也得提一下,你不能说结果得到的是一个平面图形,而是一个立体.我会向你描述它,更确切地说,是类推法会向你描述它.

“开始,我们是有一个点.当然,因为它本身只是一个点,所以只有一个端点.

“一个点产生一条线,它有两个端点.

“一条线产生一个四边形,它有四个端点.

“现在你能回答你自己的问题了:1,2 和 4,显然是几何级数.它的下一项是什么?”

我:“8.”

球:“对.这个四边形产生了一个你不知道的东西,我们称之为立方体,它有八个端点.现在你相信了吧!”

我:“这种东西有边吗?有角吗?有你们所谓的端点吗?”

球:“根据类推法来看,它当然都有.但是还得提一句,它的边可不是你们的那种边,而是我们所说的‘面’,也就是你们所说的立体.”

我:“那么,由我的身体向这个所谓‘上’的运动而产生的这个你叫做立体的玩意儿有几个立体,也就是你所说的面?”

球:“你怎么也问起来了?你不是个数学家吗?任何物体所有的‘边’—— 这里是笼统地一概这样称谓的 —— 总比具有它的物体在维数上低一个.一个点没有维,它有零个边;一条线可以说有两个边 —— 因为可以称一条线的两端为它的边;一个四边形有四个边.于是便有 0,2 和 4.你叫它什么级数?”

我:“算术级数.”

球:“下一项是几?”

我:“6.”

球:“太对了.你看你已经回答了自己的问题.由你产生的这个立方体由六个边组成,也就是说,有六个你的身体.你现在全明白了,嗯?”

“你这个大怪物!”我厉声叫道,“你这个变戏法的,弄巫术的,托噩梦的,要花活的东西!我再也不能容忍你对我的要弄了!我跟你拼个你死我活!”一边说着,我一边向他奋力冲去.

2.5 徒劳一场的球又求助于行动

真是徒劳!我用自己最坚硬的一只角猛地向这个陌生人戳去,用足以杀死一个圆的力量向他压去,但只感觉到自己根本用不上劲,因为他从我这里滑脱开了,既不是向左,也不是向右,真说不出他究竟是怎样离开的.人虽然不见了,我还能听到这个入侵者的声音.

球:“你怎么不讲理?我本希望你 —— 一个通情达理的人,一个有造诣的数学家 —— 能成为一个三维真理的合格的热心倡导者呢!对于三维世界的真理,我一千年也只能宣传一次.可现在,我真不知道怎样才能使你相信我的话.且慢,有了.我来用行动代替语言来宣示真理吧.我的朋友,请听我说.

“我已经告诉过你,从我所在的空间能够看见你们认为是密闭的一切物体





的内部.例如,在你的附近,我看见你的小橱柜里有几只你们称之为盒子的东西(像二维国里的其他东西一样,它们既没有顶也没有底),那里面放满了钱.我还看见那里面有两本账簿.我这就降下去,进小橱里去拿它一本出来.我曾看见你在半小时前锁上了那口小橱,还知道你拿着钥匙 …… 我从空间降下来了 …… 你看,这些门原封未动 …… 我已经拿到了小橱里的一个账本 …… 我又上升了 ……"

我冲到橱柜前面,使劲拉开柜门.真有一本账不见了!陌生人又带着一阵嘲笑声,在房间的另一隅露面了.这时,那个账簿也出现在地板上.我把它拣起来.不错,肯定就是刚才不见了的那一本.

我恐怖地呻吟着,怀疑自己是否失去了理智.陌生人又继续说道:"真的,你现在该明白了,我的解释是说得通的,而且也只有我的解释是说得通的吧!你们称为立体的东西,其实都只是极薄极薄的扁片;你们所谓的空间,实际上只是一个很大的平面;我在空中鸟瞰这些物体的内里,而你们只能看见它们的外皮.如果你能鼓足勇气,自己就能离开这个平面,只要稍稍升起一点或下降一点,你就会看见我所见到的一切.

"我升得越高,你的平面就离我越远,而我所能看见的就越多.当然,我所能看到的一切会越来越小.例如,我正在上升,现在我能看见你的一个六边形的邻居和他家的几个房间 …… 我看见戏院里的十扇门都打开了,观众刚刚散场出来 …… 在另一边有个圆正在他的书房里,坐在许多书边 …… 现在我要回到你这儿,而且作为一个王牌证据,我将极轻地触摸一下你的内脏,看看你会怎么想.这对你不会有任何伤害,与你将能得到的精神上的收获是不能相提并论的."

我正要提出抗议,突然感到腹内一下刺痛,好像还从我身体里发出一阵大笑.过了一会儿,这阵突如其来的疼痛消失了,只留下一点麻木感.陌生人又出现了,他渐渐变大起来,说:"我没有伤害你,是吧!如果现在还不能使你相信,我可就无计可施了.你还有什么可说的吗?"

我下定决心,再也不能让这个专横跋扈的神秘来访者肆意搞什么名堂让我的肚子发痛了,但愿我能成功地把他牢牢在墙上钉住一会儿就好了!

我又一次用我最坚硬的角向他戳去,还大声喊着,以把全家人叫醒前来相助.我相信,在我的攻击下,陌生人这下子看来很难再滑脱了.不管怎么说,他现在已不再有什么动静.我觉得仿佛听到有人前来给我帮忙了,便加倍努力地顶住他,一边继续高声求助.

一阵猛烈的颤动通过球的全身.我好像听到他说:"不,绝不能这样.要么是让他服从理性,要么是我向文明的最后成果求助."说完这句话后,他对我大声叫道:"没有必要让别人看到你所看见的一切!叫你妻子立即回去,不要让她进

来!三维的真理不会就这样失败的.不要白白地等待又一个一千年的流逝!我听到她来了.你走开,走开!快离开我,否则你就得随我一道离开这里去三维国,你的结局如何可就难说了!"

"坏蛋!疯子!不规则的东西!"我叫喊着,"我决不放开你,你该受冒名顶替的处罚!"

"呸!你竟然敢这样说!"陌生人勃然大怒,"那你送死去吧!从你的平面上消失吧!一、二、三,跟我走吧!"

2.6 我如何来到三维国,以及在这里的所见所闻

一种难以言传的恐惧攫住了我.我面前先是一团漆黑,然后又变成一片纷乱,根本不像是真用眼睛看到的景象.我看线不像线,看空间不像空间,甚至看自己都觉得不像自己了.我好不容易才能转动自己的舌头,于是便拼命叫起来:"这儿是个疯狂世界,要不然就是地狱!"

这位球用平静的声音回答道:"二者都不是.这儿是三维之国.你再睁开眼睛,平心静气地看看吧!"

我睁开了眼,于是看到了一个新世界.

出现在我面前的是一位真的可看见的圆,它比我以前靠推断和猜想捉摸出的圆更为完美.这个陌生形体的中心部分完全显露在我的眼前,但我却看不见他的心、肺和血管.我所看到的只是一个漂亮而匀称的东西 —— 你们三维国的读者们称之为球面,可在我们那里却找不出词语来称呼它.

现在,我在精神上完全被这位引路人折服了,于是喊道:"啊!您真是美和智慧的化身呀!我能看到你的内里,但那里怎么没有内脏呢?"他回答说:"你以为看到了我的内里,其实并非如此.我和二维国的国民是不同的种族,任何人也看不到我的内脏.如果我是个圆的话,你的确能看到我的肠胃,但我曾告诉过你,我是由许多圆构成的一个球体;就像立方体的一个面而是四边形一样,球的外表轮廓表现为圆."

虽然这位老师高深莫测的谈论仍令我疑惑不解,但我原来的火气再也没有了.我从内心里开始崇拜他.他用温和的口气继续说道:"即令一开始时你无法深入了解三维国的情况,也用不着苦恼.你会一步步了悟的.让我们先一起回首你的故土二维国瞥上一眼吧.我会让你看看你常常思考和推断,但却从未真正看见过的东西 —— 一个可以用眼看到的角."

"这不可能!"我叫起来.

然而,我在他的带领下梦游般地行进,直至他的声音又将我唤回现实:"看那边.那里就是你家的五边形房屋,还有住在里面的人."

我往下看去,真的看到了我的家和家中所有的人 —— 在此之前,我只能靠



推测想象到这一切呢.与我现在看到的情景相比,原来的推测和臆想是多么粗糙而朦胧啊(图22.9)!我看到在西北侧的几个房间里,我的四个儿子正安睡着;靠南墙的房间里是我那两个失去双亲的孙子.我的女儿、管家和仆人也都在各自的房间里,只有我的爱妻在为我的失踪而担心.她不在自己的房间里,而是在门厅里走来走去,焦虑地等待我的归来.刚才被我的喊声惊醒、跑到我房间里来的小听差,也正借口要弄清我是否昏倒在什么地方,向我书房的柜橱里张望呢.这一切都是我亲眼得见的,决不是凭推测得知的.我们越来越靠近我的屋子了,我甚至能看清柜橱里的两个钱匣,还有同我在一起的这位球先生曾提到的账本.

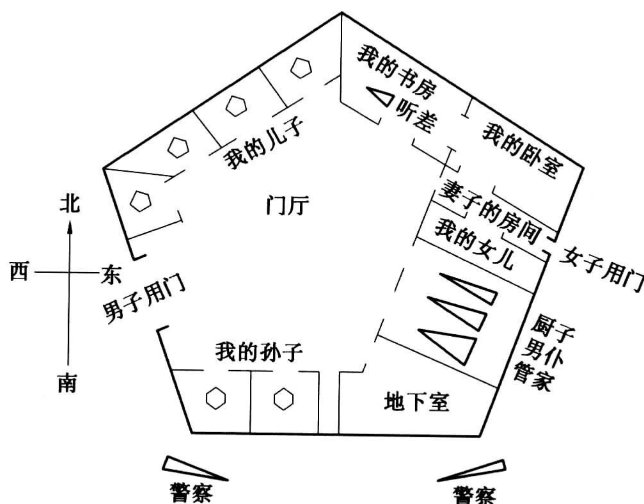


图22.9

我被妻子的悲痛所感动,想要跳下去安慰她,可却发觉自己不能动了.我的同伴说:“不要为你的妻子担心,她的忧虑不会很久,此刻让我们眺望一下整个二维国吧.”

球先生一边这样说,一边再次同我在空间升起.我的视野一步步扩大.我生活的城市,城里每座房屋的内部,以及里面的人们,这一切都尽收眼底.

我为毫无遮蔽地暴露在面前的大地的种种奥秘所震慑,便对身边的伙伴说:“我莫不是成了上帝?因为我国的圣人说过,谁能看见一切:谁就是上帝.用他的话来说,便是只有上帝才是无所不见的.”

我的老师用不无嘲讽的口气回答道:“若真是如此,那么,我们那里的扒手和强盗岂非也要被你们的圣人当做上帝来崇拜了?因为他们也能看见你现在所看见的一切呀!请相信我,你们的圣人说错了.”

我:“可是,除了上帝,又有谁能无所不见呢?”

球:“我不知道.不过,我国的每个扒手或强盗都能在你的国度上看见一切,可决不会因此而能成为上帝.你们所谓的无所不见 —— 三维国里的人并不常用这种说法 —— 能使人更正直、更仁慈,更无私和更博爱吗?一点儿也不能.那么,它怎么会使人超凡入圣呢?”

我:“‘更仁慈,更博爱’!这可是妇女的品性啊!我们只知道圆比直线更高级,原因就在于知识和智能比情感更有价值.”

球:“我可不愿把人分成三六九等.不过,在我们三维国,许多最优秀和最聪明的人认为,仁爱比知识更重要,也更看重你们歧视的直线,而并不怎么欣赏你们顶礼有加的圆.好了,不谈这个了.看那边,你知道那座建筑物吗?”

远远地,我看见一个巨大的多边形建筑群.我认出了其中的二维国议会厅.在它的周围是一排排密集的五边形建筑,建筑物之间有成直角的通道,我认出那些是街道.

我的引路人说:“我们从这里下降吧.”我便觉出自己正在靠近这个巨大的市中心建筑.此时正值凌晨,是我们纪元的第两千年的第一天第一个小时.我国最高阶层的圆们正严格地遵照古制举行隆重的秘密会议.第一千年之始和纪元零年之始时,都曾有过这样的聚会.

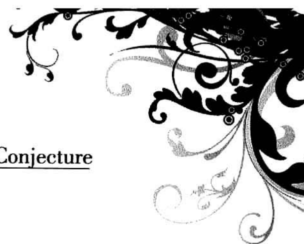
这一次的会议已经开始了几分钟,有个完全规则的四边形正在宣读上一次会议的决议.我立刻认出这就是我的在议会里当秘书的哥哥.决议上有这样的字句:“鉴于曾有图谋不轨者出现,诡称自己曾得到另外一个世界的启示,并妄言能进行诸类蛊惑人心的演示,故每逢新的一千年的第一天,将有特命官员到二维国各地组织严格的搜查,逮捕此类误入歧途者.对这类人可径直处理,无须通过数学测量.凡属等腰三角形者,无论其顶角为何均以处死论;凡为等边三角形者,一律在处以笞刑后送入监狱;四边形和五边形则交地方管教所羁留;更高阶层的成员则直接押送首都,由议会审查和判决.”

这时,议会正在第三次通过这一决议.球先生对我说道:“听到你的命运了吧?处死或监禁在等着你这个三维真理的使者呢!”我回答:“不会的.我现在明白得一清二楚,真正空间的本质是十分明显的.我想就是对一个孩子,我也能让他明白.请允许我此刻就降落下去教育他们吧!”

“现在还是别忙着下去,”我的引路人说,“这是将来的事.此刻,我必须完成我的使命.你先留在这里别动.”说着,他极敏捷地跳到二维国的茫茫人海中(如果我可以如此形容的话),而且正好落到议员中间.“我来了!”他喊道,“告诉你们,三维国的确是存在的!”

我看见,面对在他们面前不断变大的球的截面,许多年轻的议员真是惊骇不已.不过,主持会议的圆议长却了无惊慌之意.他做了一个手势,立刻便有六个最低等级的等腰三角形从不同方向向球冲去.“抓住了!”“错了,没抓住!”“没





错,抓住了!""他要跑!""他跑掉了!"他们纷乱地喊着.

议长对那些与会的众议员诸圆说道:"诸位切莫惊讶,据只供我个人查阅的一份机密档案记载,在过去的两个千年之初时,都曾发生过类似的事件.当然,你们不会在议会之外泄露这种小事情的."

他随即又提高声音对卫士们说:"逮捕这几个警察!堵上他们的嘴!你们应当知道自己的职责."处置了这些可怜的警察——不准泄露的国家机密的不幸的和不情愿的目击者——之后,圆议长又对议员们说:"诸位,会议已经结束,我谨祝你们新年好."离开前,议长又滔滔不绝地对我那位为人极好可又很不幸的哥哥说了一大堆话,说也由衷地遗憾,根据历史上的先例以及保密的需要,他不得不宣布对我哥哥实行终身监禁,还说好在用不着采取死刑——除非他再把今天的事情捅出去.

2.7 球向我展示了三维国的其他秘密,我却仍然不满足.事情的最后结局

看见可怜的哥哥被押送囹圄,我真想跳进议会厅,替他去受这无妄之灾,至少也向他话别一番.可我发现自己一动不能动——我得绝对服从我这位引路人的意志.他用郁郁不乐的口气说:"别记挂着你哥哥了,或许以后你会有充裕的时间去慰问他.现在还是跟我来吧."

我再次升入空间.球说:"迄今为止,我只向你展示了平面图形的里里外外.现在,我该向你介绍一下立体,并向你展示一下它们的构成方式.看,这儿有一些可以移动的四边形纸片,我把一片放在另一片的上面,不是像你想象的那样放在另一片的北面,而是上面.现在看,我再放一片……又放一片……看,我用这些相互平行的四边形拼成一个立方体……拼好了.它的高与它的长与宽相等,我们称它为立方体."

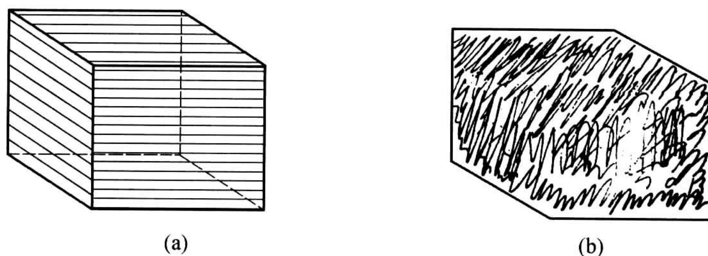


图 22.10

"请原谅,阁下,"我说,"可在我看来,它好像是一个我能看到其内部的不规则图形.换言之,我认为我所看到的不是立体,而是我们能在二维国里推断出来的平面图形.它是不规则的,这表明它是邪恶的罪犯,因此我不高兴看到它."

球说:"其实,你只是把它看成平面了,因为你不习惯于光、影和透视原理.

这正如在二维国里，一个没有视觉辨认能力的人把一个六边形看成一段直线一样. 可实际上，这是一个立体，你可以靠触摸认识到这一点.”

说着，他便领我去见识这个立方体. 我发现这个奇怪的东西果然不是平的. 他具有六个平平的表面和八个称为立体角的顶点. 我记起了球先生说过的话 —— 这样的生物可以由一个四边形在空间与其本身平行地移动而产生. 我不禁高兴地想到：即令像我这样的一个凡夫俗子，也能如此这般地缔造出杰出的后辈呢！

但是，我仍然不能完全搞清楚这位老师所说的什么光啦、影啦、透视原理啦等字眼，于是便毫不犹豫地向他说出了我的困惑.

球对我进行的解释十分透彻，不过这对于一个已经了解这一切的三维国人来说未免冗长乏味，因此我便不复述，只是告诉读者们，球先生通过他的明晰阐述，通过物体位置和亮度的变化，通过让我触摸了几个物体(甚至还让我摸了摸他本人)，终于使我明白了一切. 于是，我能一下子辨识出圆和球，以及其他平面图形和立体图形来.

这就是我多舛一生的顶峰，是我的极乐时期. 此后，我不得不叙述我的悲惨际遇了 —— 这可是十分悲惨、十分不公平的遭遇啊！难道把对知识的渴求唤醒，就是为了横遭挫折和失望之苦吗?

我不愿痛苦地追忆蒙受的耻辱. 不过，如果我能唤醒二维人或三维人的叛逆精神，使他们起来反对自以为是的二维观、三维观或任何有限维观，那我甘愿做第二个普罗米修斯，忍受目下的一切苦难，就是再不堪的境遇我也甘之如饴. 一切自私的念头，滚开！我既然干了，就要坚持到底，沿着历史的客观道路前进，既不偏离方向，也不瞻前顾后. 真正的事实和真理的声音在我的大脑里澎湃，它们迟早会一丝不差地为人们所掌握. 就请读者们来评判吧！

球还愿意继续教下去，告诉我所有规则立体的知识，像圆柱体、圆锥体、棱锥体、正五面体、正六面体、正十二面体和球体等. 可是，我还是打断了他的讲述. 不是我厌倦知识，相反，我一心想的是汲取他能提供给我的更深刻更丰富的知识.

“请原谅，” 我说，“虽说我不再称您为尽善尽美之存在，但我还是请求您，让鄙人看一眼您的内部吧.”

球：“我的什么?”

我：“您的内部：您的胃、肠子什么的.”

球：“你怎么竟不合时宜地提出了这样过分的要求！你说我不再是尽善尽美之存在，这是什么意思?”

我：“阁下，您的智慧教会了我去追求比您本人更伟大、更美好、更趋于完善的存在. 您是由许多圆结合成的，由是比二维国所有的人都优越. 因而，无疑会





有一个高于您的、由许许多多球体结合为一体的尊者存在,他会比三维国的所有立体更优秀.恰如我们此时在空间俯瞰二维国时能看见那里一切物体的内部一样,一定也存在着在我们之上的某个更高、更圣洁的区域.阁下您一定会带我去那儿的.啊!无论到达任何维的境界,我将永远称您为我的神父、我的圣人和我的朋友.在那个更广阔的空间里,在那个具有更多维的世界上,我们能从更有利的地位上,鸟瞰下面的世界.所有立体的内部,所有球体及您自己的内部,都将完全展现在我这个二维国的可怜的流亡者的视野里,那我可是锦上添花了呢!"

球:"得了!住嘴吧!别再唠叨了.时光短暂.你要想胜任向你们二维国有眼无珠的愚民们宣传三维真理,还得学许多东西咧!"

我:"不,我的好老师,您不要拒绝您有能力做到的事情吧.就让我看一眼您的内部,我将永远心满意足,从今以后永远做您温顺的小学生,做您终身的仆从,准备接受您的一切教诲,一字不漏地听您的指教."

球:"好了,我马上就回答你的请求,好让你不再啰唆:如果我能做到你所希望于我的,我是会满足你的.但是我不能.难道只为你看我的内部一眼,就把我的内脏剖出来吗?"

我:"您已经带我进入了三维国,让我看到了所有我们二维国民的内脏.因而,您现在要带着鄙人做第二次旅行,进入更高的四维境界,不是会更容易吗?在那里,我们能俯瞰三维国,看见每座三维房屋的内里,看见立体大地的秘密,看见三维国矿山的宝藏,还能看见每个立体生物,包括高贵可敬的球体的内脏."

球:"可是,哪里又有四维国呢?"

我:"我不知道.但您无疑是知道的."

球:"我不知道.没有这样的地方.这种想象简直不可理喻."

我:"阁下,这对于我都并非不可理喻,那么对于先生您就更好理解了."

"我想,即使在这儿,在这个三维国里,您也有办法让我体会到第四个维.这正如在二维国时,您能巧妙地打开您有眼无珠的仆人的眼界,让我虽没有看见,但使我确信第三维的存在一样.让我回顾一下过去吧.当我在二维国时,您不是告诉我,我所看见的一条线和推断出的一个面,实际上是具有一个我还不知道的、你们叫做'高'的,与亮度所表现的内容不同的第三个维吗?现在,在这三维世界里,我不是可以照此推想,认为由我们所看见的一个面推断出的立体,实际上也具有一个我还不认识的,虽然是无穷小的和不可测量的,然而是确实存在的第四个维吗?除此之外,用类推也可以证实这一点呢!"

球:"类推?胡扯!怎么类推?"

我:"这可是阁下您逼鄙人回顾一下您曾经揭示过的一切呢!要知道,您告

诉我的绝非是什么不足挂齿的小事.我是渴求得到更多知识的.无疑,我现在不能看见高于三维的国度,我没长着这样的眼睛.但是,正像可怜纤细的一维国王不能在二维国土上左右转动发现第二维,而二维国仍然存在一样;还正像我这个可怜的睁眼瞎不能触摸到第三个维,而第三个维却近在眼前,而且包围了我一样;四维物体也一定存在着.阁下您一定是可以用您的思想之目觉察到它的.它的存在正是您指点我的结果.难道您忘记了您教给我的东西了吗?在一维中一个点的运动,不就产生出一条具有两个端点的直线吗?

“一条线在二维中的运动,不就能产生出一个具有四个端点的四边形吗?

“一个移动的四边形在三维中的运动,不就产生出 —— 虽说我的眼睛看不出 —— 一个有幸具有八个端点的立方体么?

“那么,根据类推法,将这一真理延伸下去,一个立方体不就能在四维中运动吗?而且,运动的结果,不就应当是产生出一个比非凡的立方体更加超凡的、具有十六个端点的物体吗?

“注意这一系列有说服力的数字:2,4,8,16,不正是几何级数吗?这难道不正是 —— 请允许我引用您自己的话 ——‘完全可以类推出来’的吗?

“再有,您不是教过我,一条线有两个端点,一个四边形有四条边线,所以一个立方体一定有六个四边形的面吗?再看看这个有说服力的数字系列:2,4,6,这不正是算术级数吗?因而,由非凡的立方体通过在四维国里的运动产生的超凡结果,一定有八个立方体为其端体,这不也正像您自己对我所说过的那样,是‘完全可以类推出来’的吗?

“阁下,我的好阁下,您瞧,我其实并不真的知道什么,只是靠推测有所结论而已.我请求您,要么肯定我的逻辑推断,要么否定我的逻辑推理.如果您说我错了,我就接受您的结论,而且从此再不提什么四维;可是,如果我是对的,那您也该服从理性啊!

因此,我要问一问,在你们那里,有没有什么人曾目睹比自己更高等的生物凭空而降,进入关闭的房间,不用打开门窗,就可以随意出现和消失,正像您进入我房间时一样?对我来说,对这个问题的回答至关重要,如果回答是否定的,我以后就永不提起这个话题.务请您给我一个回答.”

球:(踌躇了一会儿)“据说有这种事.但人们对这个问题的看法是有分歧的.即使确有其事,也还有许多不同的解释.不过,虽说有许多种解释,但其中并不包括四维理论.咱们别谈这种琐节末事,还是言归正传吧.”

我:“我确信这一点.我坚定地相信我的预言会被证实.最好的老师啊,请宽容地回答我的另一个问题吧:曾经在你们国度中出现 —— 没人知道他们来自哪里 —— 又消失 —— 也没人知道他们去了何方 —— 的那些生物,在他们消失时,是否是一点点缩小其形体的截块而莫名其妙地消失在更广阔的空间里呢?”





球:(不快地)“不错,是这样消失的 —— 如果他们的确出现过的话.但是,大多数人说这些都是头脑中的幻觉引起的.也就是说,这是出现在这些所谓目击者思维中的幻象,是大脑皮层受刺激的结果.你弄不明白这是什么意思吧?”

我:“他们是这样说的吗?别相信这种话,如果肯定有这种事的话,那么这个更高级的世界就该是四维国了.那就带我去那里好了.我要在那个神妙的地方观看一切立体的内部;我会欣喜万分看到立方体在全新的方向上移动,同时严格地遵从类推法的结论行事,从而造成它的内部的每一个点都通过新的空间,形成它自己的轨迹,创造出比它自身更加至善至美的形体 —— 具有十六个超立体顶角、外面围着八个立方体的东西.而且在此之后,我们难道就会到此为止吗?在这个有四个维的神圣世界里,难道我们会踯躅在五维世界的门口徘徊不进吗?啊,不!让我们的抱负同肉体一起飞升吧.在我们智慧的冲击下,六维世界的大门将洞开,以后是七维、八维 ……”

天晓得我还能说多久.球先生大吼着命我住嘴,威胁说我要是一味说下去,他就要弄死我.可这一切全是徒劳.我如醉如痴地对真理的渴求已不可遏止地迸发出来.也许这是我的错,但又是他将我引入这一结局的.不过,这一切很快就结束了.我被什么东西从外部撞了一下,于是话头被打断了;同时又被什么东西从内部撞了一下,结果我便以高速在空间运动起来.下落!下落!下落!我迅速地降落.我知道自己在返回二维.我往下看了这沉闷的荒原一眼,这又将成为我的整个宇宙了.然后,我便陷入了一片黑暗.我的最后印象是一声巨响.当我又恢复知觉时,已经又是一个只会在平面上爬行的普普通通的四边形了.我正置身于自己的书房里,听到的是妻子走过来时发出的“轻柔的叫声”.

2.8 我在梦中受到球的鼓励

我虽然没有时间考虑,却本能地觉得必须对妻子隐瞒我的经历.这倒不是认为对她讲出来会有什么危险,而是因为我知道,这场奇遇在二维国的任何一个妇女看来都肯定是无法理解的.所以我便撒了谎,说我不小心从活板门掉进了地下室,摔晕了过去.

可是在我们这里,向南的引力相当微弱,因此就是妇女,也会觉得我叙述的经历未免离奇得难以置信.但是我妻子的脾气远远比一般女性平和.她觉察出我的不同寻常的激动,就没和我争辩.但她坚持说我病了,需要卧床休息.我倒也愿意有个借口退避到我的卧室去,好安安静静地回想发生的一切.当我终于一个人独处时,觉得自己昏昏欲睡,但在闭上眼睛前,我仍重演了一遍三维理论,特别是由一个四边形移动而形成立方体的过程,虽不像我希望的那样清晰,可我记住了它必须“向上,而不是向北”.我决定把这句话作为线索牢牢记住,相信如能牢记它,就能得出正确的结果.于是,我像念咒语一样机械地重复这句

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

话,念着念着,我便沉睡了过去.

在睡眠中,我做了个梦.我觉得我又到了球先生的身边.从他神采飞扬的样子,我知道他已不再生我的气,又变得十分温厚了.我的这位引路人让我注意一个无限小的亮点,并带我向它奔去.当我们接近它时,我好像听到了一种细微的声音,就像你们三维国的一只绿头苍蝇发出来的动静,只是更空荡,更微弱.直到我们离它不足二十个人的距离时,这声音才传到我的耳鼓里.

"看那边,"我的引路人说,"你曾在二维国住过,曾在梦中遇到过一维国,也曾同我一起飞到三维国;为了使你有完整的体验,我领你降到存在的最底层.这就是零维王国.也就是没有维数的深渊.

"看那个可怜虫.那个点也像我们一样有生命,可是被限制在无维之谷.它自己就是它的整个世界,就是它的全部宇宙;除了自己以外,它没有任何其他的概念,不知道何者为长、宽、高,因为它没有体验过;它甚至不晓得 2 这个数,他也不认为有更大的数,因为它自己是 1,1 便是一切.他其实什么也不是,却表现得十分自满.不知你注意到没有?请记住这个道理:自满是可鄙,是无知,追求要胜过盲目而无为的幸福.你听!"

他不说了.我听到这个正在接近我的小不点儿发出一种细微、低沉、单调而独特的嗞嗞声,好像是你们三维国老式留声机的动静.从这片嗞嗞声里,我听出了这句话:"存在便是无上之福.除彼之外,岂有它哉!"

"什么?"我说,"这个小东西所说的'彼'是指什么?"

球先生答道:"指它自己呀!不知你有没有注意到,有些小娃娃或孩子气的人不能把他们自己与世界区分开来,他们就是用第三人称来称呼自己的.别作声!"

"彼充满全部空间,"这小东西自言自语地继续说:"充入何处,即为何物;念及何物,即行宣布;一俟宣布,悉听不误.彼实集思者、言者、闻者为一身,亦聚思想、语言、听觉于一体;既是一,又为总.幸哉,福哉.存在便是福耶!"

"你不能让它从自鸣得意中清醒过来吗?"我说,"告诉它所处的实际地位,就像你曾告诉我的那样.向它说明零维国的狭小限度,领它升到高些的境界吧."

"这可不是一件容易的事呀!"我的引路人说,"你来试试看!"

于是,我便把嗓门提到最高,向这个点喊道:"别嘟囔了,你这可怜的角色!什么'总',实际上你什么都不是!你所谓的宇宙,其实只是直线上的一个点,而一条直线也只比 ……"

"别说了,别说了,你讲的够多了,"球先生打断了我,"现在听听你对零维国国王高谈阔论的效果吧!"

听到我刚才的话,这位国王成了更亮的光点 —— 这说明它越发志得意满





了.我刚刚住嘴,它便又拿腔作调地说了下去:"乐,思想之乐!思之所为,无坚不摧!念及彼无足轻重,愈令其幸福有加!始于甜蜜之反论,终于胜利之结局!噫!集总于一的至妙创造力!乐!存在之乐!"

"瞧,"我的引路人又说,"你的话没起什么作用呢.即令他听到了只言片语,也是把你的话当做它自己想到的东西接受的,并以此自矜'它的思想'的丰富多彩.因为它无法理解,除了它自己之外,世界上还会有别的任何事物.我们别去管这位零维国的上帝,让它自顾自地去酝酿什么无所不在和无所不知吧!无论你还是我,都无法把它从自满中解救出来."

此后,我们又慢慢地回到二维国.我的同伴和蔼地指出这番经历的寓意,支持我对知识的渴求,并鼓励我去唤起其他人对知识的向往.他承认,当初,由于我要飞升到比三维国更高的世界中去,他曾不高兴过,但后来,他有了新的认识.他并不耻于在学生面前承认自己的错误.然后,他便引导我进入比我已经了解到的更高的神奇境界,告诉我如何由立体的移动形成超立体,再由超立体的移动产生超超立体.这一切都十分简单,都是"完全可以类推出来的".

2.9　我试图向孙子传授三维理论,其过程及结果

我醒了过来,心中十分兴奋,随后便开始考虑今后应有什么大作为.我认为应马上对全体二维国民,包括妇女和士兵在内,宣讲三维真理,并决定先从我的妻子做起.

正当我打算实施这一计划时,听到街上传来人声.有不少人喊着嚷着,叫大家肃静,接着便有个更响的嗓门讲话了.这是一个传令官在宣读公告.我留心听着,听出这是议会的决定:任何人如果打算以什么别的世界的启示将人民引入歧途,便将被逮捕、监禁或处死.

我考虑了一下.这样的危险可忽视不得,最好的防范措施是只字不提我受启示的经历,只是直接进行宣讲.这样做看来十分简单而又可靠,并不会因放弃前一部分而影响实际效果."向上、而不是向北",这是全部宣讲的关键.在我睡着以前,这句话的含义有我看来是很明确的.当我第一次从梦中醒来时,也仍显得一清二楚.可不知怎么搞的,现在我竟又觉得不那么明白了.因此,尽管我的妻子恰在此刻走进了房间,在交谈了几句家常之后,我还是决定不从她开始为好.

我的几个五边形儿子都很有个性和主见.他们都是有名望的医生,可对数学却不很精通,因此不是合适的对象.于是,我便想到了我的一个孙子——一个有数学才能的年轻而温文尔雅的六边形,正是我最理想的学生.为什么不先以他为施教对象呢?他在偶然情况下想到了 3^3 的意义,这曾得到过球先生的赞许.一个毛孩子是不会知道议会决议的,因此同他进行讨论十分安全.要是同我

的儿子们谈,我可没有把握,因为他们盲目的爱国主义感情和对居统治地位的圆的盲从,可能会使他们发现我当真主张“异端” 的三维学说,于是乎自觉自愿地把我送交政府当局.

不过,我先得用某种口实来满足我妻子的好奇心.她自然希望知道有关那位“圆” 要求神秘会见的缘由,以及他进入屋子的手段.我便精心编造了一套话,不过恐怕并不像三维国读者们希望的那样能自圆其说.到底是怎样一套话,这里就不细述了,反正最终我成功地骗过了她,让她平静地将注意力转到了家庭琐事方面,不再打听什么三维世界了.随后,我便马上把我孙子找来.说实在话,我已经觉出,不知怎么搞的,我曾经见到和听到的一切,已好像一场朦胧的梦那样,逐渐变得模糊了.因此,我急于想尝试我的布道技巧,接纳我的第一个弟子.

我孙子进屋之后,我便小心地锁上门,然后坐到他旁边,让他拿出他学数学用的石板 —— 在你们看起来只是一段线 —— 告诉他继续昨天的课程.我再次告诉他一个点如何在一个维上移动而产生一条线,以及一条直线如何在两个维上移动而产生一个四边形.然后,我强忍住笑说:“喏,小淘气,你昨天要我相信,一个四边形可以按同样的方式产生别的图形,即一种三个维的超四边形,只是此时移动的方向应‘向上、而不是向北’.小嘎子,你再说一遍好不好?”

就在这时,我又听到了传令官在外边街上的喊声:“注意啦,注意啦 ……”—— 他又在公布那份议会的决议了.我的孙子虽然年轻 —— 以他的年龄论可谓聪明过人,又在尊重贵族圆的权威这样一种环境中生活 —— 但对局势的敏感是我未曾料到的.他沉默着,直到决议的最后一句话音消失,然后突然哭了起来,说道:“亲爱的爷爷,那只是我说着玩儿的,当然没有什么用意.一点儿也没有!我说这话时,咱们都还不知道这些新法令呢.我也不记得自己说过什么第三维.什么‘向上、而不是向北’ 更是没影儿的事儿.我压根就没说过.一个物体怎么会向上,而不是向北运动呢?向上而不向北?虽说我是个孩子,也不会这么胡说啊!嘻嘻!这有多傻啊!”

“一点也不傻!” 我发火了,又拿起手中的四边形对他说道,“就拿这个四边形打比方吧.我移动它.你看,不是向北,而是 …… 对,向上 …… 也就是说,不是向北!而是向别的一个什么方向 …… 不完全是这样,而是以某种方式 ……” 我的话就这样卡住了,再不知道怎样说下去,只是漫无目的地挥舞着手中的四边形,把我孙子逗得大笑不止.我从来没有听到他这么响亮地笑过.他说我不是在教他,而是在逗他玩.说着,他便打开房门跑了出去.我给一个小学生上的第一堂三维真理课便告完结.

2.10 我试图以其他方式散播三维理论,结局如何

在孙子身上的失败,倒是没有促使我向家庭的其他成员泄露我的秘密,但





也没有令我灰心.这次尝试使我看出,单靠“向上、而不是向北”这句话是不行的,应当努力寻求一种论述方式,以把全部有关观点清楚地摆在公众面前.因此我觉得应付诸写作.

所以,我一连几个月闭房不出,专心撰文论述神秘的三维世界.只是为了不冒犯法律,我不得不避而不讲具体的维.我说的是一个思维国.从理论上说,从这个国度里能俯瞰二维国,看见一切物体的内部和外部,而且在那个国度里,可以存在一种形体,它的外面被六个四边形和八个顶点包围着.但在我写这本书时,发现最糟糕的是不能根据需要,画出必要的图解来.因为在二维国里,人们能看到的并不是图形而只是线.一切都表现为直线,只有长短和明暗的不同,因此这篇文章(我给它起名为《从二维国到思维国》)完稿时,我拿不准究竟会有多少人明白我的意思.

与此同时,我的生活中出现了阴云.我对一切娱乐都感到厌倦,眼前的一切都在劝诱我说出反叛的话来.因为我不能不把我在二维国看到的情况与它们在三维国里的实际面貌相比较,而且简直要忍不住大声说出来.我轻慢了我的当事人,忽视了自己的职责,终日沉思在我曾看见,但不能告诉任何人的神秘事物中.此外我又发现,我头脑中的印象已越来越难复现了.

在我从三维国回来后的第十一个月后的一天,我闭上眼睛,试想在大脑里想象出一个立方体,然而却失败了.后来我又成功了,但已不那么把握十足,而且以后也一直如此.这使我更加消沉了.我决定采取行动;但采取什么行动,我又说不出.如果能够说服众人,我是宁愿献出生命的,可是如果我甚至连亲孙子都说不服,又怎么能让我国最高级和最进化的圆们信服呢?

有许多次,我在精神极度亢奋的状态下冒出了危险的言论.当局已经认为若不是心怀叵测,那就是迷上了异端邪说.对此,我是一清二楚的.可是,我仍不时管不住自己,说出令人生疑和离经叛道的话来.我甚至把这类话说给最高级的多边形和圆听.例如,当一次讨论对自称能看见物体内部的狂人的处理意见时,我就引用了古代一位圆的话说:预言家和天才总是被多数人视为狂人.有时,我偶尔还会说“能看到物体内部的眼睛”和“能一览无余的地方”.有一两次,我甚至脱口说出了“第三维和第四维”等被禁言语.最后,我从这些“小意思”竟发展到了大不韪,在地方官员本人的官邸举行的一次“地区思辨学会”聚会上放了一炮.在这次聚会上,我听到一个愚不可及的糊涂虫读了一篇洋洋洒洒的文章,大谈其上帝把维数限制为两个的理由,以及为什么只有上帝才能无所不见,便忘形起来,详细地叙述了自己的经历:我随着一位球体进入三维空间,进入了市中心的议会厅,然后再次进入空间,最后又返回家中…… 我讲了我实际上和在梦中看到和听到的一切.一开始时,我还假借了一个虚构的人物来描述一段假想的经历,可我的激情使我很快抛弃了一切掩饰.最后,在这个热

情洋溢的长篇演说的结尾，我呼吁在场的所有听众抛开偏见而皈依三维真理.

结果还用说吗?我立刻遭到逮捕.

第二天早上，我被带到了议会受讯问.我恰恰站到了几个月前球先生同我在一起的位置上.他们让我讲述我的故事，没有人提问，也没有人打断我.可是从一开始，我就预见到了自己的命运.因为在我进行辩护前，在场的警卫都是些顶角较大的等边三角形，顶角都在55° 上下.而到了我发言时，议长却命令把他们撤了下来，换成了顶角为2° 和3° 的更低等卫士.我心里雪亮，知道这意味着我很可能被处死或监禁，听到这一叙述的所有在场者也将被一道消灭，以向社会保密，因此议长要用低劣者替较有价值者送命.

在结束我的辩护之后，也许是觉察到一些资历较浅的圆们已被我的坦诚所感动，议长问了我两个问题：

(1) 我能否指出我所说的“向上、而不是向北”这句话所意指的具体方向?

(2) 我是否能用图解或叙述的方法(而不是类推什么假想的边和角)表明我所言及的立方体图形?

我声明再没有什么可说的了，只说我必须尊重真理，而真理终将胜利.

议长说他很赞成我的这句话，还说比这更好的结束语怕是不会有了.他判我无期徒刑，还说如果果真有什么三维真理认为我应当隔离缧绁、出来向世界宣讲真理的话，那么就让我等待着这些真理亲自来解救我好了.不过他又说，除了要采取一些防范我逃跑的必要措施外，我在牢狱中无须吃苦;而且只要我不利用这一优渥待遇胡作非为，我还能见到先于我入囹圄的哥哥.

七年过去了.我依然是个囚犯，而且除了看守我的狱卒和偶然来看我的哥哥外，再见不到任何人.我的哥哥是一个极好的四边形，他正直、聪明、乐观，又富手足之情.但我承认，我同他每周一次的晤面，至少在一个方面使我更加痛苦:当初球先生在议会厅出现时，我哥哥也在场，他看见了球的截面的变化，也听到了球对圆们所做的表演和讲解.在七年的牢狱生活中，几乎没有一个星期我不同他反复讲述我在球先生那次演示中干了些什么，我还大量地叙述了三维国的种种现象，以及由类推法推断立体的存在.但是我不得不承认，他仍不能领悟第三维的本质，始终不相信球这种形体的存在.

因此，我是一无所成.从我的所有经历来看，千年一度的启示并没有使我做出任何成绩.三维国里曾有位普罗米修斯，他把天火带给人类，结果为此而受苦.而我这个可怜的二维国的普罗米修斯，虽然也在受苦，却什么也没有给我的人民带来.不过，我仍抱着希望，希望我的回忆能以某种方式 —— 什么方式，我可想不出来 —— 传到什么地方，激励那里的人们奋起，不再囿于有限个维的世界.

这就是我心绪较好时的希望.不过，我并不是总能如此.我时时会心情沉





重,觉得我不能毫无愧怍地宣称对自己所见过的,并往往为此抱憾的立方体的确切形状一清二楚.在夜里,我总是在梦中听到"向上、而不是向北"这句话,就像是中了邪.我固然有为真理献身的勇气,但也会不时陷入软弱的精神状态——当立方体和球体在我脑海中显得很不真实时,当三维国看起来也同一维国或零维国一样缥缈时,当限制我自由的坚硬墙壁、我用来写字的便笺,以及二维国的种种实在在我面前像是痴想的产物和荒诞的梦幻时.

生活空间的维度[①][②]

第二十三章

1 维度数学

只是在近年里,数学才找到了满意地处理维度问题的方法.维度差异不是大小差异或空间潜力差异.人们可以把一条直线的点集同等于一个有界两维区域或一个三维物体的各点.在考虑多少维度时,人们应该将其归因于生活空间,不必考虑描述时空间"可用的"纯数量问题.

数学表明,维度是空间的"内部结构"的一种属性,这种属性密切联系于拓扑属性.两维空间例如平面的属性是,在平面之内,不可能把5个或5个以上的点互相连接而连线不相交.而且,我们提到,在两维空间中,不可能把圆内一点和圆外一点连接而同圆的边界不相交.在三维空间中,这类道路则是可能的.这些事实清楚地表明,心理生活空间的维度数目是多么重要,例如对于那些位移可能进行的问题.

在现代维度理论中,维度数目的确定,本质上取决于空间物体和它的边界之间的一般关系.例如,一个三维空间物体的边界是面,即这种边界本身是两维的.一个面的边界是线,即一

① 原载:库尔特·勒温.拓扑心理学[M].竺培梁,译.杭州:浙江教育出版社,1997.

② 库尔特·勒温(Kurt Lewin)是德国心理学家、德籍犹太人.1890年9月出生于普鲁士的莫吉尔诺(今在波兰),先后求学于弗赖堡、慕尼黑、柏林等大学.1914年在柏林大学获得哲学博士学位.除学习心理学之外,他还学习数学和物理学,1922年勒温任柏林大学讲师,1927年升任教授,1932年赴美讲学,在斯坦福大学任教.1933年他定居美国,在康奈尔大学任教,1935年转至依阿华大学任教授,1944年转至麻省理工学院,主持团体动力学研究中心的工作,并担任加利福尼亚大学伯克利分校和哈佛大学的访问教授,1947年2月去世.





个两维空间的边界是一维的线.一维有限线的边界是本身零维的点.人们可以说:一个 n 维空间的边界,是$(n-1)$维的.很清楚,如果不详细研究关于这些问题的数学,而根据这个事实,递推程序也是可能的.人们能够开始于把空集作为(-1)维空间,从它继续进行到较高的维度.

从数学的观点来看,没有理由把维度数目限于3个.发展到 n 维空间可能毫无困难.数学也处理在不同点上维度数目不同的空间.也许看来迷人的是,心理学有可能自由使用较为复杂的空间或多维空间.然而,出于科学和经济的原因,提出理论时,引入不多于绝对必需的维度,这在开始时是十分重要的.因此,我们的问题必须表述如下:描述生活空间所需要的维度的最小数目是多少?

也许应该把我们的问题区别于另一个问题,在那个问题中,也可能在数学意义上讲到维度.人们能够认为一个物体或一个系统的不同属性是变量,并能够用坐标系来描述变量,这种坐标系的维度,同可区分的属性一样多.物理学在这种意义上讲到相空间(用同样的方法,能够把时间描述为第四维度).类似地,波林(Boring)讲到心理学中刺激的维度.很清楚,生活空间的维度数目的问题,完全不同于心理物体或事件的可区分属性的数目的问题.

2 心理环境的维度

现实性 在我们实验工作的起初,两维空间足可描述情境.然而,后来我们便使用三维空间了.

在确定生活空间的维度数目时,参考位移概念和上面提到的区域维度以及区域边界之间的关系,这是有益的.生活空间的维度数目大于零,这是由于运动不可能在零维空间进行.位移的事实意味着生活空间至少是一维的.

一维空间的边界本身是零维,它们相当于点.点是心理区域边界的一种合适描述吗?把生活空间描述为一维的即一组线,乍一想,它看来同事实相矛盾.然而,用数学上无可争辩的方法来否决这种描述,倒并不容易.因为一维空间能够由十分复杂的网络所组成.例如,人们能够如此来定义一维空间,以至于有可能在每一点上区别任何数目的方向.然而,我认为,实际发生的心理过程的性质,使得把生活空间描述为一维的是不可能的.在心理域内,"沿着边界"走,这时常是可能的.而且,人们通常能够把一个连通边界的组分区分为不同的扇形.我们提到,边界的不同组分在它们的强度上可能不同,这对于个体行为是重要的.这些事实表明,存在着不是由点组成,而是至少由线即一维流形所组成的边界.至少具有一维的边界,意味着我们在讨论至少是两维的空间.

人们是否应该使用更多的维度,回答这个问题较为困难.当我们讨论准物理域时,也许有必要把心理域描述为三维空间,它相当于物理空间的3个维数.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

例如,要围绕一个两维障碍,某些位移在3维物理区域的点之间是可能的,而在两维空间内则是不可能的.根据这些位移的属性,人们也许能够表明,像物理空间一样,准物理域也具有3个维度.

非现实性的程度 对于我们目前讨论的大多数问题来说,准社会域和准物理域的两维描述就够了.然而,当我们必须区分不同的"现实性程度"时,就有必要过渡到更多维度.白日梦,也就是模糊希望,一般来说,它的现实性比行为少;行为的现实性有时比言语多;知觉的现实性比想象多;遥远的"理想目标"的现实性,比决定人们即刻行为的"现实目标"少.行为本身有时具有不同程度的现实性.凡是和个体的迫切需要有关的过程,通常具有高度的现实性,在这种过程中,个体必须克服坚固的物理障碍或社会障碍.例如,在准概念过程之间,人们能够把深思熟虑达到某一目标的方法,区别于较为非现实性的自由幻想.

把心理学中现实性程度的概念,区别于认识论上现实性概念,这是十分重要的.对于认识论概念来说,没有现实性的不同程度.存在和不存在之间的中间体是不可能有的.心理学上"非现实"物体,对于认识论意义上的心理学是现实的,即这些物体存在,它们产生效果.必须进一步强调,有关物体的物理现实性,并不决定心理现实性的程度.对于原始人或生活在"奇幻世界"的儿童,他们所相信的那些事实就是现实的.

实验表明,现实性程度几乎是所有心理物体和心理过程的一种十分重要的动力属性.这点尤其表现于志向水平、替代行为的起因和效果、目标的形成和变化、情绪过程、记忆、游戏等实验中.

我们发现,有必要使用一个特殊维度来描述现实性程度.如果人们把现实区域总体描述为两维空间,例如一个平面,那么人们必须把现实性的不同程度同等于一个在另一个之上的不同平面.为了表明这种描述在数学上的必要性,人们将必须考虑到,位移可能在现实性的同样程度之内进行,也可能在现实性的不同程度的区域之间进行.属于现实性同样程度的总体,例如个体在某一时刻的希望世界,本身是一个区域.在这个区域内,人们能够区分组分区域.我们能够讲到在非现实层面内的运动,如同在现实层面内一样.因此,有可能确定这些区域的拓扑学.一个非现实性层面的拓扑结构有时类似于现实性层面.然而,在某些情况下,尤其当现实性层面内的情境令个体十分讨厌时,非现实性层面的结构一般不同于现实性层面的结构.

为了遵循使用不多于必需的维度来描述生活空间这条原则,人们必然会问,是否可能描述非现实性水平而不引入一个新维度.缜密的考察表明,这类描述是行不通的.

如果人们试图描述现实性程度较低的区域而不使用另外一个维度,人们就想起有点相同的两种方法.人们可能试图把逐次减少的现实性的水平,描述为



区域链,每一区域相应的现实性程度要低于它前面的区域(图 23.1(a)).或者,人们也可能把它们描述为一系列环状区域,每一区域完全被后面区域包围(图 23.1(b)).

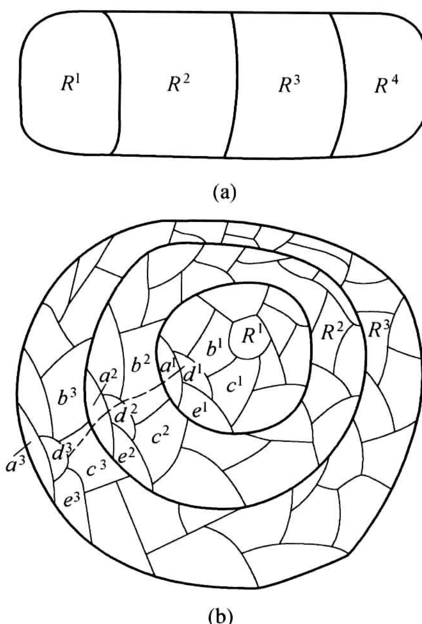


图 23.1　试图描述现实性不同程度而不引入另外的维度.现实性的不同程度在(a)中被描述为区域链,在(b)中被描述为一系列环状区域 R^1, R^2, R^3, R^4($R^1 \supset a^1$, $R^1 \supset b^1$, $R^1 \supset c^1$, $R^1 \supset d^1$; $R^2 \supset a^2$, $R^2 \supset b^2$, $R^2 \supset c^2$, $R^2 \supset d^2$; $R^3 \supset a^3$, $R^3 \supset b^3$, $R^3 \supset c^3$, $R^3 \supset d^3$);区域 d^1, d^2, d^3 的内容相同

然而,这类方法不适于描述某些位移.如果人们在现实性的同样水平之内,区分每个组分区域,这点就变得清楚了. $a^1, b^1, c^1, d^1, e^1 \cdots, d^3, e^3$ 可能是现实性程度 R^1 的组分区域; $a^2, b^2, c^2, d^2, e^2 \cdots$ 可能是现实性程度 R^2 的组分区域; $a^3, b^3, c^3, d^3, e^3 \cdots$ 可能是现实性程度 R^3 的组分区域.现实性程度的差异可能不是很大,现实性的不同水平的结构大致相似.而且, R^1 中的区域 d^1, R^2 中的区域 d^2,和 R^3 中的区域 d^3,就它们的有关内容而言,可能在心理学上属于同类.

两维描述有着不足之处,现实性的同样水平之内的运动,例如从 b^1 到 a^1,几乎总是意味着接近或离开现实性程度较高或较低的区域.而且,困难还在于,如果人们描述从一个区域到另一个内容相同但现实性程度较低区域的位移,那么它首先必须越过现实性程度相同的其他区域.例如,为了从 d^1 运动到 d^2 和

d^3,人们必须越过区域 a^1, b^2, d^2, a^2, c^3.在实际的心理生活空间之内,人们也不能直接从一种现实性程度的每一个区域,运动到另一种现实性程度的每一个区域.但是,在相同内容的一个区域内,当然会发生越过一系列不同现实性程度的心理位移.如果人们把不同现实性程度的水平描述为一个两维空间,就不能把这些位移描述为连续道路而不歪曲不同水平内区域的连通性关系.这点表明,如果人们要合适地描述这类位移,就必须引入一个新维度来描述现实性程度的差异.

顺便提一句,这些考虑也使维度的基本数学属性更为清楚.只有当人们破坏拓扑关系时,才能够使不同数目维度的空间内的点或其他组分区域一一对应地同等.两维描述(图 23.1(b))破坏 d^1, d^2, d^3 之间实际存在的关系.

由于不少于两维的空间相当于同样现实性程度的区域,因此如果人们考虑现实性的不同程度,则至少必须用 3 个维度来描述生活空间(图 23.2).我们把不同现实性的程度同等于一个在另一个之上的平面或水平.在图示中,我们将用位于较高的平面来描述非现实性(Ⅰ)程度较高的水平.当然,人们也能够进行相反描述.

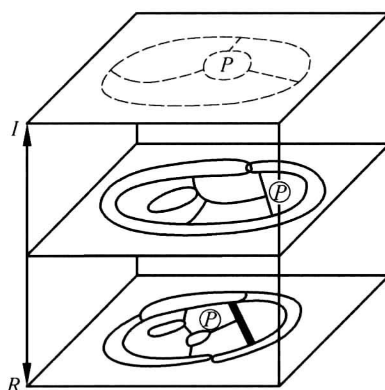


图 23.2　由生活空间的另外一个维度来描述现实性的不同程度 R,较为现实的水平;I,较为非现实的水平;P,个体.在较大现实性水平上,障碍较强,个体 P 同他的环境分离较为清楚

现实性不同程度之间的最重要动力差异之一,便是较为非现实的水平(Ⅰ)有较大的流动性.下列不同事实中表明这种较大的流动性:(1) 环境内的障碍产生的阻力相对很小(人们能够在非现实性中为所欲为);(2) 环境区域的边界之改变较为容易,而边界之确定则较不明确;对于现实性不同程度的准概念区域,情况也是如此;(3) 内部一个体的张力系统的扩散释放较为迅速地发生,这





相当于内部系统的较弱的墙;(4) 个体和环境之间的边界较不清楚,环境的结构在较大程度上取决于个体的需要.

3 个体维度的问题

人们提出问题,用于描述生活空间内个体的区域,具有多少维度.仅仅在现实性平面内描述个体,这将是错误的.因为个体也能够在较为非现实水平内行动.而且,看来存在着人们必须把它们同等于非现实水平的内部一个体系统.

这个问题引出另一些问题,而对于那些问题,目前几乎不可能作出明确详细的回答.毫无疑问,人们必须把解决一个问题中从认真推理到幻想尝试的过渡,描述为到非现实水平之内过程的过渡.但是,在这些情况下,我们在讨论整个个体的位移吗?假如个体在每一时刻都处于现实性的唯一一种水平之内,人们就必须把个体描述为两维的.然而,这种假设是令人怀疑的,因为甚至在跃入非现实性期间,个体至少部分停留在现实性水平之内,例如他的身体.这点由下列事实证明,即使个体进入幻想,他仍能够受到在现实性水平之内过程的影响.可能发生的情况是,这种影响的可能性大为减弱,以至于人们必须把个体描述为动力上相对分离的两个区域,它们属于现实性的不同水平.在精神分裂症中,有时可能看到这类状况.

人们也可能想到把个体描述为一个区域,它同时属于现实性的一种以上水平.在这种情况下,像环境一样,必须把个体描述为至少三个维度的区域.它可能时而较远、时而较近地进入现实性的不同平面,个体的"重心"(center of gravity) 可能在不同时刻属于不同水平.

最后,在大多数情况下,可以认为向现实性的较低程度水平的过渡,是个体组分的一种位移.这类个体组分区域的位移,并不必然导致个体统一性的破坏.毫无疑问,我们在现实性水平之内发现这类运动.

整体上,我似乎认为,心理学事实表明,个体和整个生活空间的维度数目相同,可用有关的生活空间所具有的现实性的所有水平来描述个体.把个体向现实性的其他水平的过渡,描述为现实性的不同水平的相对权重的一种变化,这通常是可能的.例如,有可能把个体 P 同时所处的现实性的不同水平 R_1, R_2, $R_3\cdots$ 处理为一种交叠情境.在一定时刻,整个个体的行为,在不同程度上受到现实性不同水平的影响(R_1 的相对权重 W_R 可能是 70%,R_2 的相对权重可能是 5%,R_3 的相对权重可能是 10%,等等).不同水平的这些相对权重,在不同情况下以及对于不同个体,都会有所不同."跃入非现实性"将意味着,较高现实性的某些水平的相对权重(例如 W_{R_1})减少了,而起初可能为零或仅仅稍大于零的生活空间的某些较为非现实水平之权重(例如 W_{R_4})得到相当增加(例如,起初

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

$W_{R_4}=5\%$，现在 $W_{R_4}=40\%$）."跃入非现实性"这种描述避免了许多困难，而用个体从现实性的一种水平到另一种水平的位移来描述，自然碰到这种困难. 它具有的另一优点是，重视在个体内不同区域的重要性的变化，这种变化是这类过渡的属性. 当然，这种变化不仅是结构变化，而且也是功能变化.

在高度现实性的水平内，区分个体和环境是相对容易的（我们在这里不考虑边界确定的准确性问题）. 无论如何，并不难以表明，环境中存在着个体没有占据的地方. 然而，在高度非现实性的水平内，人们通常不能确定，应该把某一区域描述为个体的组分，还是环境的组分.

个体和环境分层为现实性的不同程度的水平，这不应该混淆于个体内中心层和边缘层之间的区分. 因为在现实性的同样水平之内，人们将必须区分较为中心的内部区域和较为边缘的内部区域. 整体上，在较为非现实的平面内的过程，看来与个体的核心区域和他的中心需要，具有更为密切的动力关系. 人们是否必须仅仅在现实性的水平内来描述个体的运动层面，这是另一个特殊问题.

4 生活空间在现实性－非现实性维度上的分化

不只当幻想尤为活跃时，生活空间才分层为现实性的不同程度的区域. 生活空间总是如此. 然而，这种分层的程度取决于整体情境. 有时候，个体向现实性的不同程度水平的过渡，是生活空间（即个体和环境）在现实性－非现实性维度上的扩张或收缩，而不是在一个常定区域内的位移. 而且，现实性的不同程度水平的相对权重能够发生变化.

生活空间朝非现实性方向扩张多远，这对于不同个体，对于不同时刻，看来是十分不同的. 人们可能问道，在这个方向上是否存在着生活空间的内在限度，即是否存在着一个最高的非现实性的平面. 现在几乎不可能证实这种假设.

倒很可能证实最高现实性水平的假设. 人们可能注意到运动层面同现实性的一种特殊关系，从而试图建立在每一生活空间内存在着这种"完全"现实性的平面这个论题. 另外，我们也有可能在概念上把完全现实性，联系于从"外面"对生活空间的客观的物理的或社会的影响. 就有关的动力学而言，现实性的本质属性是，不以个体意志为转移. 然而，也有一些事实违反这种完全现实性的绝对水平的假设. 有人可能相信，他生活在完全现实性中，直到他认识到更加确凿的现实性的事件为止.

如果人们想到生活空间从儿童期到成年期的发展，那么，关于最高现实性平面的相对性的事实，就变得更为重要. 我们已经提到，这种发展的特性是个体和环境的扩张，尤其是分化. 在现实性差异的维度上，人们也能够观察到生活空间的这类逐渐分化. 大量心理事实证实这些假设：(1) 在儿童的生活空间内，现



实性的程度上只有些微差异(图 23.3(a));(2) 儿童生活空间内的各种水平,整体上相当于现实性的中级程度(图 23.3(a) 和 23.3(b)).这意味着儿童现实性的水平的属性为相对非现实,与成人相比,这也意味着儿童非现实性的水平相对现实.因此,看来我们不必引入"绝对现实性"平面这个概念.我们已经提到,对于不同的个体,对于不同时刻的同一个体,生活空间分化为现实性的不同水平之程度各不相同.在人们能够明确回答这些问题之前,进一步的调查研究是必需的.但是我认为,为了在生活空间的拓扑学之内描述非现实性的水平,我们所讨论的概念和方法,足可解决任何概念上的困难.

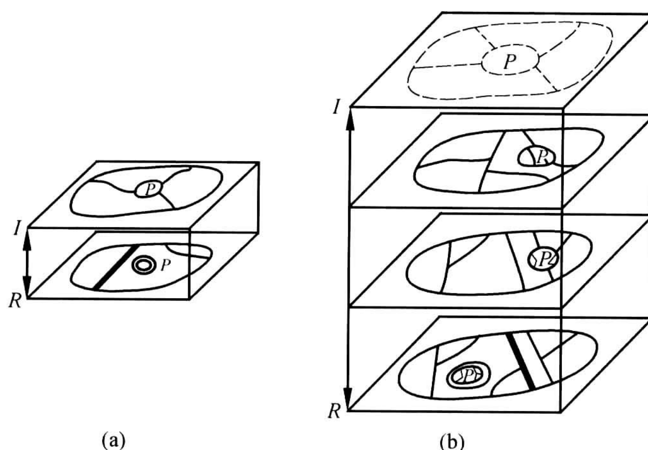


图 23.3 成人的生活空间在现实性 - 非现实性维度上分化较高.儿童生活空间内现实性层面和非现实性层面的范围,相当于成人生活空间内现实性层面和非现实性层面的中间范围

我们在本书中所发展的各个概念,涉及整个心理生活空间,即个体和环境.这些概念准许人们论述关于生活空间及其组分的位置和联系的所有问题.它们适用于准物理的事实,也适用于准社会的事实和准概念的事实.运用这些概念,人们能够描述个体和环境的结构变化,以及各种各样的位移.在一定程度上,人们也能够论述那些狭义的心理动力问题,例如区域的摩擦力、与位移有关的障碍的固体性、与不同类型的交往有关的区域分离的程度、环境系统和个体系统的统一性(动力格式塔)程度等.最后,人们能够论述关于区域状态的张力和变化的某些问题,例如区域的流动性或固体性.

但是,这些动力问题以这样或那样的方式,联系于超出拓扑心理学的种种问题.拓扑心理学仅仅能够描述生活空间内"可能"事件的框架.为了确定实际上发生哪些事件和这些事件的冲突基础,人们必须考虑有向的量,尤其是心理力的概念和力域的概念.我们把有关的概念和同等定义,作为向量心理学的组分.

附　　录

附录Ⅰ　庞加莱,杰出的数学家、科学家和科学哲学家

——纪念庞加莱逝世八十周年①

庞加莱(Poincaré,Jules Henri)②被公认是19世纪最后20年和20世纪初的一位杰出的数学家.他1854年4月29日生于法国的南锡,1912年7月17日在巴黎逝世,终年只有58岁.在他并不太长的一生中,不但在数学的众多领域都有开创性的重大成就,而且对物理学和天体力学也有卓越的贡献,同时,他还是一位很有影响的科学哲学家.在现代科学家中,庞加莱也是一位值得重视和认真予以研究的人物.本文概述庞加莱的生平、成就和思想,以表示对他逝世80周年的纪念.

1

自19世纪的后半叶起,整个欧洲形成帝国争霸的局面,同时,无产阶级革命运动崛起,先后成立第一国际、巴黎公社和第二国际,并出版了马克思的巨著《资本论》.那时的法国不但外患经年,国内更处于政治动荡之中,第二共和国变为拿破仑称帝的第二帝国,直到1871年第三共和国成立,其间出现过多起重大的政治事件.不过在1870年普法战争之前,也有过一段相对安定的时期,法国的工业革命正是在这时兴起.社会的动荡,促使人民思想活跃,激起许多人去深刻地思考各种社会问题.在法国涌现出雨果、左拉、福楼拜、莫泊桑、玻特莱尔等一批大文豪,形成了法国文学的黄金时代.一批印象派画家也于这时出现.

正是在这一时期,一系列的革命性的重大科学成就相继问世:达尔文进化论,麦克斯韦电磁理论,弗洛伊德心理学,电子、X射线和放射性元素的发现,以及量子论和相对论的提出等等.这些科学新成就向人们的传统观念发出了强烈的挑战,使人们的思想逐渐革新.同时,在工程技术方面则有发电机、电灯、无线

---

① 原作者孙小礼.

② 中译名曾有普恩加莱、彭加勒、庞卡莱、邦加来等,本文采用中国大百科全书数学卷的译名庞加莱.





电报等一系列新发明,直接地改变着人们的生活方式.

庞加莱正是成长和生活在这样一个社会动荡不安、思想空前丰富多彩,科学文化空前发达的时代.

庞加莱的祖父和父亲都是医生.祖父曾在军队中行医,父亲的医德更为人称道,晚年曾任医学院的教授.他的叔父是桥梁道路督察,堂兄弟雷蒙曾是法国总理,并在第一次世界大战期间荣任法国总统.庞加莱自幼天资聪颖,受到家人宠爱,母亲对他尤其爱护备至.可惜他五岁时染上了白喉,九个月不能说话.这场病使他成为一个瘦弱怕羞的孩子,总是回避同那些粗野的男孩一块玩耍.然而他识字很快,能够独自看书,爱不释卷,而且过目不忘.六七岁时读过一本《洪水前的大地》,产生极大兴趣.他的老师,一位优秀的小学教师,善于满足他的好奇心,诱导他提出种种问题,使小庞加莱不但学得很多知识,还养成了勤于思考的习惯.九岁时,庞加莱的第一篇作文就名列全班第一,同时也逐渐显露出非凡的数学才华.

1870 年普法战争爆发,南锡市首当其冲,遍地伤员、惨不忍睹.16 岁的庞加莱成为一名救护员,随父亲在战地医院工作,救护了很多伤病员.这一经历使他亲身体验到战争烽火之苦楚.

在 1871 ~ 1872 年间,法国公立中学举行数学联赛,庞加莱获得了第一名.在面试中,老师向他提问一道几何题.庞加莱先在黑板上粗粗地画出图形,思索了一会就回答道:“只要对两条线用极反演立刻可以得到答案.”老师说:“非常有理,但我想要一个更为简单的答案.”庞加莱在黑板前徘徊了一会,突然有所发现,他说:“用三角学方法可以求得答案,-….”对于几何题一般都在几何范围内寻求答案,而庞加莱却能想到用三角学来解答几何题,使老师感到异常满意.在南锡,庞加莱成为一名不同凡响的学生,在他的数学考卷上虽然会有演算错误,或者没有答完试题等情况,但是他那种出色地思考问题和提出问题的方式却总是博得老师们的赞赏.据说南锡的一位教授曾把庞加莱称为“数学巨怪”(Monster of Mathematics).但是,庞加莱的绘画能力很差,同学们开他玩笑,把他的绘画搜集起来开展览会,对每幅画都要用文字说明,“这是一匹马”或是“这是一幢房子”.然而庞加莱成为数学家以后,欣赏绘画、音乐、艺术却是他的爱好,而且认为数学家应当像艺术家那样地工作.

1873 年 10 月,庞加莱以优异的成绩考入巴黎综合工科学校.这所由法国数学家蒙日于 1794 年创办的理工科大学,当时已是法国的科学研究中心.庞加莱于 1875 年毕业,数学成绩最为突出.

1875 ~ 1878 年,庞加莱又进入巴黎高等矿业学校学习,毕业后被政府聘任为主管 Vesoul 区的矿务工程师.有一次发生瓦斯爆炸事故,十六人罹难,庞加莱不顾自身安危,亲自下井视察,表现出对工作的高度负责精神,博得矿工们的爱

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

戴.而在这一段时间里,庞加莱的个人志趣仍然是钻研数学问题.他于1878年写出一篇讨论微分方程的解的性质的论文,正是这篇论文使他于1879年获得巴黎大学数学博士学位.

1879年,庞加莱到卡昂大学数学系任教,从此开始了他的数学创造生涯.1881年起他成为巴黎大学教授,直到1912年去世.1885年庞加莱荣获巴黎科学院的Poncelet奖.1887年他被选为法国科学院院士.1908年被选为法兰西学院院士,这是当时法国学者的最高荣誉.

庞加莱每年讲授一门不同的课程,有纯数学、应用数学和物理学的各个分支,以及天体力学方面的课程,都受到学生的欢迎.他的讲稿一般都很快被整理成书出版.庞加莱在数学、物理学和天体力学方面的专著和教科书多达三十部之多.

庞加莱习惯于来回踱步时在脑子里思索数学问题,并且进行心算和推理,当他思考成熟后很快地写下来,基本上不用修改或重写,就能形成论文.他一生中发表专业论文近五百篇.人们公认他是一位勤奋多产的科学家.

庞加莱在科学界成名之后,声望日益增长,人们邀请他在许多公众场合、对数学和科学的各种问题作演讲,他乐意接受这类邀请,总是热情地、认真地向人们通俗地解释数学和科学中的概念和方法,宣传数学和科学的意义和重要性,而且由于他的思路清晰、语言生动,特别引人入胜.他的名著《科学与假设》、《科学的价值》、《科学与方法》,以及《最后的一些想法》都被认为是科学哲学方面的著作.因为庞加莱的文章和著作,文笔流畅、风趣,使人像读散文一样,所以他被誉为法国第一流的散文大师,还被选为法国文学会会员.庞加莱曾拥有多种头衔,是四十多个法国的和外国的学院和学会的委员,前后接受过剑桥、牛津、格拉斯哥、布鲁塞尔、斯德哥尔摩、柏林等大学所授予的博士学位.

庞加莱很重视数学史和科学史的研究.他曾指出:"如果我们想要预见数学的未来,适当的途径就是研究这门科学的历史和现状."1912年萨顿创办Isis杂志时,曾请庞加莱担任第一任编委会委员,可惜他不久便离开了人世.

庞加莱在自己的著作中曾详细地记述了他在自守函数研究中的创造过程,使人们能够知道这位天才数学家是怎样进行创造性的研究工作的.为庞加莱撰写悼文的茹尔丹先生在悼词中写道:"他之所以会永远活在人们心里,有许多原因,其中一个原因是他既能使我们钦佩他,又能使我们了解他."

2

庞加莱出生于高斯(Gauss, Carl Friedrich,1777—1855)逝世的第二年.人们常把庞加莱与高斯相提并论,都被看做是最高意义上的广博数学家.他们两人





不但都在数学的各个领域有主要成就,而且都对天文学和数学物理作出卓越贡献.一般认为,19 世纪的数学,起初是在高斯的直接影响下发展的,而到世纪之末则由庞加莱所主宰.庞加莱因此而被认为是那一时期的"领袖数学家".

由于庞加莱对他那个时代的全部数学都能融会贯通地加以掌握,整个数学都是他的研究领域,所以他也被称为数学的"多面手"或"全能数学家".然而现代数学的飞速发展,一个人同时通晓数学的若干领域已极不容易,所以人们又认为,庞加莱也许是达到这样一种全能境地的最后一位数学大师了.

首创精神在庞加莱的全部科学研究活动中都有充分的体现,他为数学的发展开拓出许多新的天地.但他不爱在细节上继续花费功夫,常常把细节问题留待后继者们去深入细致地加工和完善,他自己则很快又跨进新的领域了.对于他的这种风格,数学家们说,"他像一位征服者,而不是殖民地的统治者."

微分方程的理论及其对于动力学的应用,在庞加莱的数学研究中处于中心地位.从他 1878 年的第一篇论文到 1912 年的最后一篇论文,可以说几乎都是从各种可能的角度在探讨微分方程及其有关问题.差不多每一年他都发表了有关微分方程这一主题的论文.看得出来,庞加莱的许多数学创造是由各种微分方程问题引发而致.

庞加莱 1878 年论述微分方程的博士论文引导他进而研究自守(Automorphic)函数,自守函数在微分方程理论中扮演一个重要角色.庞加莱是从富克斯(Fuchs,I.L.)的一篇研究 $y''+P(z)y'+Q(z)y=0$(P,Q 是复变数 z 的有理函数)的文章出发的,发现这种方程的两个线性无关的解的商的反函数是一类新的自守函数,即指那些在包括 $z'=\frac{az+b}{cz+d}$(a,b,c,d 是实数或复数,且 $ad-bc=1$)等不连续群的变换下不变的函数.这是对三角函数、指数函数(单周期)以及椭圆函数(双周期)等周期函数的一个重要推广.庞加莱把他新发现的自守函数命名为富克斯(Fuchs)函数,而把他紧接着又发现的一类自守函数叫做克莱因(Kline)函数,Fuchs 和 Kline 都是研究自守函数的先行者.后来庞加莱指出,对于仅有正则奇点的以代数函数为系数的 n 阶微分方程,都可用他所发现的自守函数来求解.在 1881~1884 年间,庞加莱在《数学学报》上集中地发表了五篇关于自守函数的论文,自此之后,自守函数便成为函数论中的一个专门论题或分支了,而庞加莱则由于这一重要成就而一举成名,不到三十岁就进入世界著名数学家的行列.

创建微分方程定性理论是庞加莱最杰出的贡献之一.从 17 世纪后半叶起的二百年间,常微分方程的发展始终围绕着如何求解这一中心问题.人们总在追求用积分法求出用初等函数或用初等函数的积分表示的解.虽然数学家们积累了很多求解的经验和方法,但也逐渐清醒地认识到,能够求得初等函数解的微分方程只占极少数.正像在代数中那样,自阿贝尔(Abel,1802—1829)证明了

五次以上的代数方程一般没有根式解以后,人们放弃了一味寻求根式解的企图,转而探讨代数方程的系数与根的关系,使得一些数学家想到可以从微分方程本身去探讨解的性质.然而第一个明确地提出这一思想,具体地开辟出一条研究途径的正是庞加莱.他以“关于由微分方程所定义的曲线的研究报告”为题目,于1881,1882,1885,1886年连续发表了四篇长达二百多页的研究论文.他把微分方程求解的老课题转换成研究由微分方程定义的积分曲线这样一个新课题.也就是说,在不求出解的情况下,通过直接考察微分方程的系数和方程本身的结构去研究解的性质,分析和推断积分曲线可能具有的各种特性,如曲线的形状、结构和趋势等等.

庞加莱的研究工作是从一类形式简单而在实际问题中经常遇到的非线性方程:$\frac{dx}{X}=\frac{dy}{Y}$开始的.他发现使$X$和$Y$都等于零的点,即方程的奇点,对于研究解的性质具有关键性的作用.他将奇点分为四类:节点、鞍点、中心点和焦点,详细地讨论了解在四类奇点附近的结构和性态,引进了极限环等新概念.方程有极限环,就相当于有一个孤立的周期解.庞加莱把自己这一整套的研究成果称为微分方程的定性理论.定性理论的创立是微分方程发展史上的一个很重要的里程碑.正如前面所说,庞加莱把定性理论的某些细节留待后人加以完臻.瑞典数学家本狄克生(Bendixson,1861—1935)于1901年发表论文补充了庞加莱的研究结果,从而有著名的Poincaré-Bendixson定理.

庞加莱所开创的微分方程定性理论,直到现在仍然是吸引许多数学家从事研究的一个活跃领域.1900年德国著名数学家希尔伯特(Hilbert,1862—1943)作了展望20世纪数学的报告,提出了23个重要数学问题,其中第16个问题就是求一阶一次微分方程$\frac{dy}{dx}=\frac{Y}{X}$($X,Y$是$x,y$的$n$次多项式)的庞加莱极限环的个数的上界和相对位置问题.近些年来许多数学家(包括我国数学家)作了很多研究,取得了重要结果.

庞加莱是代数拓扑学的奠基人,他的这一开创性研究对现代数学也有深远的影响.事实上,他在微分方程定性理论中关于积分曲线的性状与奇点的讨论,就是一种拓扑学的研究,以后他进一步对高维空间中的曲线与曲面作了定性研究.法国数学家迪厄多内(Dieudonné)指出,“连续性”概念是庞加莱数学研究的重要动力,他在研究问题时,总要探讨一下当这个问题的条件允许连续变化时会发生什么情况,因而他都要遇到拓扑学问题.庞加莱自己也于1901年说过:“我所着手探讨的每一个问题都把我引至位置分析(Analysis Situs)”.位置分析是拓扑学的前称.庞加莱于1895年所写的“Analysis Situs”是一篇一百多页的重要论著,正是开创代数拓扑学的头一块丰碑.他自1894年起的十年之内所发表的六篇论文构成了代数拓扑学的基本内容和方法.他引进了贝蒂数、挠系数等





概念,创造了流形的三角剖分、单纯复合形、对偶复合形、复合形的关联系数矩阵等一系列新概念,以及由这矩阵计算贝蒂数的方法.借助于这些工具,他推广了欧拉的多面体定理,称为欧拉-庞加莱公式,还发现了流形的同调对偶定理.有数学家正确地指出,直到1933年发现高阶同伦群之前,代数拓扑学的发展完全是以庞加莱的思想和技巧为基础的.

1904年庞加莱猜想:单连通的三维闭流形必与三维球面 S^3 同胚.这就是著名的庞加莱猜想,是拓扑学中的重要猜想之一.后人接着猜测:当维数 $n\geqslant4$ 时,单连通的闭流形如果与 n 维球面 S^n 有相同的同调群,亦必与 S^n 同胚,这称为 n 维的庞加莱猜想.1981年M·H·弗里德曼证明了四维的情形,然而庞加莱关于三维的猜想则至今未得解决.庞加莱猜想是考验我们对于流形的认识深度的一块试金石,对这一猜想的每前进一步都能引起拓扑学的一次跃进.四维庞加莱猜想的解决导致一个非常重要的发现,即四维的欧氏空间与其余维数的欧氏空间不同,除了通常的微分结构以外,还有别的不寻常的微分结构.

在函数论方面,庞加莱是多复变解析函数论的先驱者之一.他的这一开创性工作是与研究多变数偏微分方程的解的性质相联系的.

庞加莱在代数学方面也有重要的成果.例如,他在1903年关于线性微分方程的代数积分的论文中研究了非交换代数.他引进了群代数并证明其分解定理,是他首先引进了左理想和右理想概念.庞加莱还是当时能够理解和赞赏李(Lie,1840—1899)及其后继者关于"连续群"工作的少数数学家之一.庞加莱-伯克霍夫-维特定理是李代数理论中的一个基本定理.

在数论方面,庞加莱在1901年发表的题目为"有理数域上的代数几何学"的论文,是很有影响的,开创了丢番图方程的有理解研究.

庞加莱在自守函数理论的研究中用到了非欧几何.于是,他也独立地给出了双曲几何的一种模型,后来称为庞加莱模型.这是他对非欧几何的一份贡献.

在天体力学方面,庞加莱的成就是很突出的.1885年瑞典国王奥斯卡60寿辰时,计划颁发一项科学奖金,委任柏林学院的魏尔斯特拉斯(Weirstrass,1815—1897)、法兰西学院的厄米特(Hermit,1822—1901)和斯德哥尔摩大学的Mittagliffler教授组成一个委员会,由他们提出了一个天体力学难题:根据牛顿力学规律,要求求出任意多个互相吸引的天体的运动规律.这是一个特定的微分方程组求解问题,简称"n 体问题".庞加莱写了应征论文,研究了 $n=3$,并且其中两个的质量比另一个小得多时的三体问题的周期解,还证明了这种限制性三体问题(日、地、月系统便属于这类三体问题)的周期解的数目同连续统的势一样大.委员会认为庞加莱的这一成果为天体力学的研究打开了新局面而决定把奖金颁发给他.法国政府为此给庞加莱加封勋位,使他成为法国头号学者之一.

事实上,庞加莱从1883年起就开始研究三体问题了,到1891年止,他发表

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

了九篇关于三体问题的文章.庞加莱在天体力学的大量研究中引进了新概念、新方法和新技巧,创建了动力系统理论,这些成果都集中于他的名著《天体力学的新方法》(第一、二、三卷,分别于1891,1893,1899年出版),这是一部与拉普拉斯1825年出版的《天体力学》相媲美的重要著作.

庞加莱还进行过天体演化的研究,出版了一本教材:《天体演化假说》.他在1885年发表过论文,讨论在引力作用下,凝聚在一起的旋转流体的平衡外形,他提出一种新的可能结构——梨状结构.这一假说在近些年来重新引起人们的重视,因为地球物理学家借助于人造卫星发现地球的外形可能是呈梨状的.

庞加莱发表了近百篇论文讨论太阳系理论的各种问题.迪厄多内认为庞加莱是"牛顿之后在天体力学中做出最出色工作的数学家".

庞加莱也是一位卓有成就的物理学家,尤其擅长于研究理论物理.因为他善于把数学研究和物理问题结合起来,所以在数学物理方面曾有一系列的重要成果.他在巴黎大学讲授各种物理学课程达二十年以上,出版了《流体力学》、《热力学》、《数理光学》、《塑料理论》、《电与光》,等著作.他在物理学方面的文章和著作共有七十多种.在世纪之交的物理学革命时期,他也作出了独特的贡献.

1895年底,伦琴把第一张X射线照片寄给了庞加莱,引起了他的极大兴趣.他的关于X射线与荧光现象有关的猜测,使柏克勒耳受到启发,随即进行荧光物质是否发出X射线的一系列实验,进而导致放射性的发现.

庞加莱被认为是相对论的先驱.他在1898年的题为"时间的测量"的论文中,提出了光速在真空中不变的公设.他还讨论过利用光信号使时钟同步的问题.他在自己的论文中以及在1904年在美国所作的"物理学的现状与未来"的演讲中都曾相当完整地表述了相对性原理.虽然他没有像爱因斯坦那样进一步建立起相对论这样的新理论,但是学界公认庞加莱独立于爱因斯坦,得到了狭义相对论的许多结果.

1911年12月9日,亦即庞加莱逝世前的半年,他在一封写给Rondo Circolo Meterna Tico Palermo(可能是意大利的一个数学联谊会.——原作者注)的创始人和会长Guccia的信中写道:"像我这样年纪的人,既已略有成果,正好可以带领后起的研究人员迈向新路;展望未来,实在不忍草草放弃往日的成绩.既然如此,不知你认为是否可以发表一篇未完成的论文?论文叙述我所追求的目标,我所提出的问题,以及我追求结论所用的方法.这样做看来有点不近情理,但我认为会有一定的影响力."Guccia立即发表了这篇未完成的论文,题目是"一个几何定理".这是庞加莱提出的一个拓扑定理,也是他的最后一篇研究微分方程的论文.他把有限制的三体问题的周期解的存在问题归结为平面连续变换在一定条件下的不动点的存在问题.在庞加莱逝世后不久,美国青年数学家柏克霍夫于1913年发表论文证明了这个定理.





庞加莱以微分方程为中心,触类旁通,在数学的众多领域作出划时代贡献的事实,说明数学各分支具有内在的联系和统一性,这一点在数学和科学都飞速发展着的今天也是颇有启发意义的,做科学的多面手是可能的和必要的.

3

自 1902 年起,庞加莱开始从事科学哲学方面的著述,相继完成了《科学与假设》(1903),《科学的价值》(1905)和《科学与方法》(1908)三部名著.在他逝世的第二年(1913)又出版了汇集九篇演讲和文章的文集《最后的一些思想》.他的三部名著曾被译成多种文字:德文、英文、俄文、西班牙文、匈牙利文、瑞典文、日文和中文,在全世界广为流传,尤以 1913 年在美国纽约出版的《科学的基础》,即三部名著的合编本影响最大,其中还包括庞加莱为《科学与假设》英译本所写的序言.在我国,商务印书馆曾先后出版了中译本:《科学的价值》(文元模译,1928)、《科学与假设)(叶蕴理译,1930 年初版,1957 年修订重印)、《科学与方法》(郑太朴译,1933).光明日报社出版社又于 1988 年出版了《科学的基础》的中译本,书名改为《科学的价值》(李醒民译,石雷校).

庞加莱的哲学思想在相当大的程度上承袭了法国实证主义思想传统.他认为科学并不能告诉我们事物的真实本性,而是告诉我们事物之间的真实关系.庞加莱更直接地受到法国物理学家迪昂(Duhem, 1861—1916)和奥地利物理学家马赫(Mach, 1838—1916)的哲学观点的影响.他很推崇马赫提出的思维经济原则,在他的著作中多次援引马赫的这一原则.他赞同迪昂的整体论观点,认为一个包含着多个假设的理论被某一个实验所否证时只能知道这些假设不同时为真,但不能判定哪一个假设为假,所以,所谓"判决性实验"是不存在的.在哲学史上,人们常把以孔德(Comte, Anguste, 1798—1857)为代表的实证哲学称为第一代实证主义,把马赫、迪昂和庞加莱等科学家的哲学思想称为第二代实证主义,而把 20 世纪 20 年代以后在西欧形成的逻辑实证主义,亦即维也纳学派,称为第三代实证主义.

著名物理学家杨振宁前几年在一次与青年的座谈中,对于物理学与哲学的关系表达了这样一种观点,即认为哲学有两种涵义,第一种是哲学家的哲学,第二种是对物理问题的中距离的甚至是短距离的看法.这第二种哲学对于物理学有着关键性、长期性的影响,因为它决定了一个物理学家喜欢提什么问题,喜欢了解什么问题,一个问题来了喜欢用什么方法去解决.这种哲学与一个人的风格、喜好有极为密切的关系,对一个人的研究成就有决定性的影响.每一个科学

家根据他过去的经验都会形成他自己的这种哲学.①

庞加莱的哲学属于杨振宁先生所说的第二种哲学.我之所以持这种看法,是以庞加莱的三部科学哲学著作的具体内容为根据的.

读了这三本书,可以感受到庞加莱撰写的科学哲学的主要内容是对他所熟悉的长期从事研究的数学、物理学、天文学这三个领域中的一些问题的哲学思考,尤其侧重于对科学方法的思考,大都是对这些问题的近距离或中距离的看法,虽然在书中也有一些"远距离"的看法,然而分量极少,不占主导地位.庞加莱的哲学思想是对他在这三个领域中的科研实践经验的直接概括和总结.

在《科学与假设》一书中,庞加莱深有体会地说明了假设(Hypothesis)的作用.他认为,凡是推广都是一种假设,因此假设有着必不可少的作用,这是谁也不会反对的.不过,它应当总是尽可能早地、尽可能经常地受到检验.当然,如果它经不住这种检验,人们就应当把它抛弃.然而物理学家的假设不是轻易采用的,这个假设已顾及到一切似乎能够参与现象的各已知因素.如果检验不支持它,其中必有未曾料到的、不寻常的东西,而这正是我们要去寻觅的未知的和新颖的东西.所以,被抛弃的假说远非毫无成效,正是给科学家提供了作出新发现的机会.可以说,它比真实的假设功劳更大.

庞加莱指出,科学家不能满足于一堆实验事实,应当做整理工作.人们靠事实建设科学,正如用砖石筑成房屋;然而许多事实不能称其为科学,正如一堆砖石不能称其房屋.最要紧的,要有预见.而要预见,就必须有推广.实验只给我们一定数目的孤立的点,我们用一条连续的线把这些点连接起来,这就是真正的推广.实际上我们画出来的曲线是在这些点附近穿绕而过的,并非恰恰通过这些点.这样,我们还对实验进行了矫正.依靠推广,一个观察到的事实可以使我们预见许多事实.当然这种预见被实验推翻的可能性很大.但是,与其根本不去预见,还不如去作不够准确的预见.

庞加莱把假设分为三类:第一类是自然的(Natural)假设.例如人们假定十分遥远的物体的影响可以忽略不计,微小的运动遵循线性定律,等等.第二类是中性的(Neutral)假设.例如,假设物质是连续的,或者相反,物质是由原子构成的.相反的假设并不改变计算的结果.这种假设的用处或在于作为计算的技巧,或有助于理解具体的图象,或坚定某种思想.第三类是真实的(Real)推广.这是由实验证实或否证的假设,不管是被论证了还是推翻了,它们都可能是丰产的.

对于任一事实都有无穷的方式进行推广,因而有一个选择问题.庞加莱说,每一种推广都在某种程度上含有对自然界的统一性和简单性的信念.选择是受简单性的考虑导引的.他进而指出,考察科学的历史,有两种相反的现象.有时

① 见《科技日报》1987年1月25日.





简单性隐藏在复杂的外表下面，有时简单性是表面的，而它隐藏着非常复杂的实在．如果我们的研究方法日渐精密，我们便可在复杂的里面找出简单的东西，再从简单里面找出更复杂的东西．如是循环不已．然而，我们必须停留在某个地方，为了科学可能起见，当我们找到简单性时，就应该停下来．这里正是我们唯一的地基，在这上面建立我们推广的大厦．

在《科学与方法》一书中，有一章专讲事实的选择．庞加莱把简单与美联系起来，他说：正因为简单是美的，正因为壮观是美的，所以我们宁可寻求简单的事实、壮观的事实．他总结说："像对于实用的渴望一样，对于美的渴望也导致我们作相同的选择．因此，按照马赫的看法，这种思维之经济、劳力之经济是科学的永恒的趋势，同时也是美的源泉和实际利益的源泉．"

庞加莱对数学推理的性质和特点作了很详细的探讨．在《科学与假设》一书中，他指出，数学推理并非只是演绎的，三段论法不能告诉我们本质上新颖的东西；它在某种程度上带有归纳推理的性质，正因为这样它才非常富有成效．他说明了递归推理是使我们从有穷通向无穷的工具．他认为数学家是"通过构造"而工作的，"构造"出越来越复杂的组合，再用分析的方法从这些组合、这些集合返回到它们的初始元素，他们得知这些元素的关系，并由此推导这些集合本身间的关系．他特别强调说："我们只有借助数学归纳法才能上进，只有它能告诉我们新的东西．如果没有那种在某些方面有别于物理归纳法但又同样有效的数学归纳法的帮助，构造则无力去创造科学．"

在《科学的价值》一书中，庞加莱专用一章讨论数学中的直觉和逻辑．在数学家中，一些人只承认逻辑，另一些人则诉诸直觉，他认为，对于科学的进步来说，这两种精神都是必要的．例如，在欧几里得创造的几何学这个庞大的建筑物中，它的每一个部件不管怎样都归因于直觉，可是，我们今天也可毫不费力地从中辨认出一位逻辑主义者的工作．他指出直觉和逻辑各自的缺陷，前者不能给我们以严格性或可靠性，后者则不能使我们创造任何新东西．庞加莱的结论是：逻辑和直觉各有其必要的作用，二者缺一不可．唯有逻辑能给我们以可靠性，它是证明的工具；而直觉则是发明的工具．

在《科学与方法》中，庞加莱解释了什么是数学创造．他认为创造并不在于用已知的数学实体作出新的组合，这样作出的组合可有无限多，大多数没有用处．创造恰恰在于不作无用的组合，而作有用的、为数极少的组合．发明就是识别、选择．数学的美感、数和形的和谐感、几何学的雅致感，正是这种特殊的审美感，在大量组合中起着微妙的筛选作用．直觉使我们推测隐藏的和谐和关系．只有某些组合是和谐的，从而同时也是有用的和美的．他的体会是"有用的组合恰恰是最美的组合，我意指最能使这种特殊情感着迷的组合，所有的数学家都知道这种情感．"

庞加莱还在“数学创造”这一章中详尽地回忆和记述了他如何写出第一篇关于自守函数的论文的经过.他把自己的创造过程概括为“两种自我的工作方法”,即有意识的自我和无意识的自我在轮番地工作着.无意识的自我决不劣于有意识的自我,它机智、敏锐,知道如何选择,如何凭直觉推测,它在有意识的自我失败了的地方获得成功.以一种突如其来的灵感或顿悟形式呈现在脑海中的组合,常常是有用的、多产的组合.当然,无意识工作之有成果是有条件的,即一方面要有有意识的工作在它之前,另一方面要有有意识的工作尾随其后.庞加莱的记述和分析是一份提供给心理学家研究创造性思维的极为宝贵的素材,至今仍是有重要价值的科学文献.

庞加莱根据切身的体验透彻地讲明了数学与物理的关系.一方面,数学为物理学提供表述定律的唯一语言,帮助物理学家进行有效的计算;最为重要的是,在从实验得到定律,亦即进行推广和选择的过程中,必须借助于数学的理性精神才能揭示隐藏在事物中的奥秘.另一方面,数学也受赐予物理学,物理学是研究自然的,认识自然的欲望对于数学的发展具有最持久、最有效的影响.“人的想象可以变化,但是自然的变化更加丰富多彩.”傅里叶级数理论、二阶偏微分方程理论等都主要是借助于物理学和为了物理学而得以发展的.他还指出,数学与物理学的证明方法不同,但是发明方法却是十分类似的.

以上我所转述的虽然偏重于庞加莱关于科学方法方面的探讨,然而是有代表性的.由此可以看出,他所表述的看法都是与数学、物理学的具体内容密切联系的一些“近距离”的哲学见解.

在20世纪初关于集合论的悖论以及数学基础的争论中,庞加莱对罗素(Russel,1872—1970)的逻辑主义和希尔伯特的形式主义都持批评的态度.(见《科学与方法》中有关逻辑的三章)正是他的这种立场使得他成为数学哲学直觉主义派的先驱之一.这部分内容或许可算是他对数学问题的“中距离”看法.

在哲学史上庞加莱还被视为约定主义的代表人物.其约定主义思想直接来源于19世纪出现的非欧几何.由于欧氏几何中的第五公设:在给定直线外的一点只能作一条平行线,分别被替换为“在给定直线外的一点至少可作两条平行线”和“在给定直线外的一点不能作出平行线”,从而构造出两种新的几何体系.庞加莱由此认为几何学的公理既非先验综合判断,亦非经验的事实,而是一种约定或公约.如果把这种观点推广到一般,则成为哲学上的约定主义(Conventionalism).然而他在《科学与假设》书中明确说明,“实验是真理的唯一源泉”.在物理学中存在着定律和原理,定律是实验结果的概括,原理才是约定的公设,而且这种约定的和普遍的原理是实验的和特殊的原理的自然而直接的推广.他批评说“有些哲学家推广得太过份了;他们以为原理就是整个科学,从而认为全部科学都是约定的.”在《科学的价值》一书的引言中他指出“一些人把约定在科学





中的作用夸大了，他们甚至走得如此之远，以致说定律，即科学事实本身是科学家创造的.……不，科学定律不是人为的创造，……”后来，有些哲学家也认为庞加莱并非真正的约定主义者，毋宁说他是约定主义与经验主义的综合.

迪厄多内在他为庞加莱写的传记中说“是否能够像通常地理解那样授予他‘哲学家’的称号是个问题，这个问题最好留给专业的哲学家去判定.”看来，和一些著名的专业哲学家相比，庞加莱并没有构建出任何哲学体系，也没有多少惊世之哲言.他的哲学思想不属于杨振宁所说的第一种哲学，充其量他只是一位科学哲学家.他的哲学著作充满着具体的科学内容，他是以科学家的身份对各种科学问题进行了哲学的思考，其看法不一定都对，有时还流露出错误的唯心主义思想.他的一些观点曾经被许多人所赞同引用，同时也遭受到多方面的批评，包括列宁的严厉批评，这在历史上是很自然的事情，都有具体的历史背景和原因.然而，瑕不掩瑜，批评并不足以抹杀其哲学思想中正确、积极的东西.

总之，庞加莱的科学哲学著作毕竟是留给后人的宝贵的精神遗产，在科学界和科学哲学界有着实际的影响.例如，爱因斯坦在 1921 年发表的“几何学与经验”一文中，称庞加莱是“敏锐的、深刻的思想家”；在 1952 年发表的文章“青年时代二三事”中说，庞加莱与马赫对他的发展有一定的影响.庞加莱在科学方法论方面有许多朴实的真知灼见，现代科学哲学各流派的不少见解都可以在他的思想中找到雏型.

附录 Ⅱ　预备知识

1　第一个存在定理[①]

这一节专用来提出本文的主要存在定理.在第 2 ~ 7 节中给出它的证明,第 8 节中进行总结.在本节中从考察一些特例着手,然后引出这定理的陈述.要记住,我们的问题是要提出一个判别准则,使能在许多情形下运用它来断定:能不能把形如 $f(x)=y$ 的方程解出 x 来.为着要看出如何提出判别准则,我们先考察一些我们已经知道如何求解的方程.

先看 x 的值在 -1 与 $+2$ 之间由公式 x^2+1 定义的函数 $f(x)$.(公式对于 x 值在区间 -1 到 2 之外时有意义,不过我们不考虑这一事实)图 1.1 表示这个函数的图形.方程 $y=x^2+1$ 定义一条抛物线,而我们的图形是介于两条铅垂线 $x=-1$ 与 $x=2$ 之间的那段抛物线.

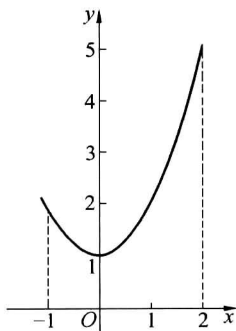


图 1.1

首先注意到曲线在 $x=0,y=1$ 处有一个最低点.再说得准确些:对于所有在 -1 与 2 之间的 x,x^2+1 大于或等于 1,而在 $x=0$ 处取极小值 1.如果要找当 x 在 -1 与 2 之间时,曲线的最高点,那就要问"在 -1 与 2 之间"的这一区间是否不包含 $x=2$ 这端点.如果包含这端点,那么最高点就在 $x=2$ 与 $y=5$ 处.如果不包含,那么曲线就没有最高点;因为不论我们在曲线上取哪一点,它的横坐标总比 2 小,在横坐标更接近于 2 的点处曲线上有更高的点.为着避免这样的情况,我们应该使区间包含端点 -1 与 2.于是,对于 $-1\leqslant x\leqslant 2$ 的所有 x 值,x^2+1 比 5 小或等于 5,而函数在 $x=2$ 时有极大值 5.

现在考虑解方程 $x^2+1=y$ 这个问题.设 y 值已经给定,求满足这方程的在区间 -1 与 2 之间的相应 x 值.若 y 值超过极大值 5,当然没有解.在 y 比极小值

① 第 1 节到第 12 节为一维时的存在定理.



1 小的情况也一样.然而,若 y 在 1 与 5 之间,就有解.这只要在 x 轴的上方作高等于 y 值的水平线就能从图形上看出来.若所作的水平线太高或太低,它就不会与曲线相交.当高在 2 与 5 之间时,它与曲线相交一次;而在 1 与 2 之间时,它与曲线就相交两次.(用 y 的式子来表示 x 时,一个式子是 $x = \sqrt{y-1}$)

作为第二个例子,令 $f(x)$ 由公式 $\frac{2x}{x^2+1}$ 对于 $-3 \leqslant x \leqslant 3$ 的所有 x 值定义.它的图形见图 1.2.从方程

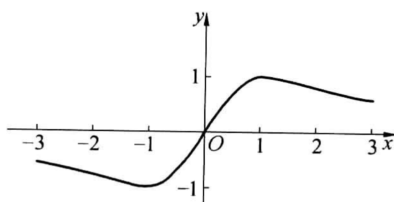


图 1.2

$$y = \frac{2x}{x^2+1}$$

不难看出:一个正 x 值给出一个正 y 值,而一个负 x 值给出一个负 y 值.还有,改变 x 值的符号只改变 y 值的符号;因此,曲线关于原点对称.曲线的最高点在 $x = 1$ 时出现,这时 $y = 1$.要证明确是如此,我们对于 $1 - f(x)$ 这个差作一点儿代数替换

$$1 - f(x) = 1 - \frac{2x}{x^2+1} = \frac{x^2+1-2x}{x^2+1} = \frac{(x-1)^2}{x^2+1}$$

因为最后的式子永远不能是负值,从而 $1 - f(x) \geqslant 0$,于是 $f(x) \leqslant 1$.由于对称性,曲线的最低点在 $x = -1$ 与 $y = -1$ 处.因而,若 $y > 1$ 或 $y < -1$,则方程 $f(x) = y$ 显然无解;但是对于 $-1 \leqslant y \leqslant 1$ 的每一个 y 值,方程能解出来.(方程的两端同乘以 $x^2 + 1$,再从所得的二次式解出 $x = \frac{1+\sqrt{1-y^2}}{y}$)

人们会从上面讨论的两个例子进行推广,并且猜测:若任一函数 $f(x)$ 对于区间 $a \leqslant x \leqslant b$ 的每一个 x 值有定义,则 $f(x)$ 有一极大值 M,一极小值 m,并且对于区间 $m \leqslant y \leqslant M$ 的每一个 y 值,方程 $f(x) = y$ 有一解.让我们多作几个函数的图形来检验这个猜测;这里要记住,我们能用图形来定义函数.

若函数 $f(x)$ 的图形如图 1.3 所示,是光滑曲线,猜测看来是对的.当 y 在 m 与 M 之间时,高为 y 的水平线必与曲线相交.即使曲线如图 1.4 所示,有些拐角,猜测仍然正确.但是,若曲线如图 1.5 所示,有一处断裂,猜测就错了,因为通过断裂处的水平线不与图形相交.图形有断裂处的这种函数,在数学中是很自然的.它们叫做不连续函数.一个不连续函数的图形甚至于可以没有最高点也没有最低点,例如图 1.6 所示的图形,它在 $x = c$ 处断裂,它的点 $(c, f(c))$ 是在 r 处.

现在我们已经做完准备,可以叙述第一个主要定理.

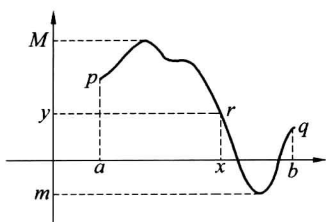


图 1.3

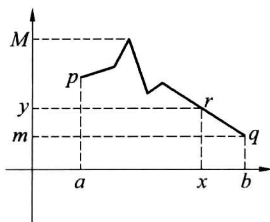


图 1.4

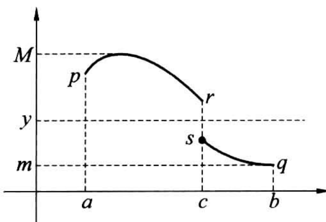


图 1.5

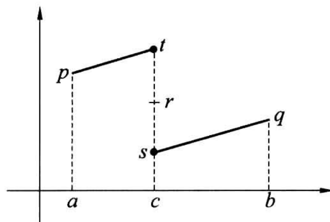


图 1.6

定理 若对于某闭区间$[a,b]$的所有实数x,函数$f(x)$有定义,取实数值,并且连续,则函数有极小值m与极大值M,并且对于闭区间$[m,M]$中的每个y值,方程$f(x)=y$在区间$[a,b]$中至少有一个解x.

定理有时简述如下:若实值函数$f(x)$在区间$a\leqslant x\leqslant b$中有定义且连续,则该函数有一极小值,一极大值,并取得其间的每一个值.

“闭区间”指的是包含端点a,b的区间,就是说,对于x值的限制是$a\leqslant x\leqslant b$.“开区间”不包含端点.用$[a,b]$表示闭区间,用(a,b)表示开区间.“半开区间”包含一个端点但不包含另一个,如$(a,b]$指的是$a<x\leqslant b$,而$[a,b)$指的是$a\leqslant x<b$.

已知一函数$f(x)$连续并给定一个特定值y.在这情形下,这定理怎样帮助我们决定能不能解出$f(x)=y$呢?若我们能确知$f(x)$的极小值m与极大值M,只要问:是否$m\leqslant y\leqslant M$?m与M在许多情况下不易找出.但是,算出函数的一些值通常还是容易的.若对于某个x值c,得$f(c)<y$,而对于另一个值d,得$y<f(d)$,则定理断言在$[c,d]$(或$[d,c]$若$d<c$)中有一个x值使$f(x)=y$.例如,若$f(x)$是$x^3-\sqrt{1+4x}$,则$f(0)=-1$,而$f(2)=5$.所以$x^3-\sqrt{1+4x}=2$在区间$[0,2]$中有一个解.

必须着重指出:这个定理的重要在于它的普遍性.在大量的千变万化的情况下,它告诉我们一个永远可靠的求解的方法.像x^2+1这样不计其数的特例,就用不着这个定理,因为我们要找的可以很容易找到.应用于复杂的函数,定理就显出它的威力了.但更重要的,它是连续函数的一般理论中的第一个定理.

还必须着重指出:这定理的目前提法不够完全.f的连续性的确切意义没



有定义;我们只给出基于几何图形的一个直觉描述 ——f 的图形是一条没有断裂的曲线 —— 但这仅是用一个无定义的术语来替代另一个.下两节要引导我们来得到连续性的确切定义.

2 集合与函数

本节从头到尾首要关心的将是几何图形.这些图形就是欧几里得直线的、平面的、或空间的子集.为方便起见,我们假设已在直线、平面、或空间中引进了笛卡儿坐标,每个点 x 用它的坐标表出.一个点的坐标形成一组有顺序的 n 个实数$(x_1,x_2,\cdots,x_n)$,简称为实数的一个 n 组.于是在平面的情形 n 是2,在三维空间,n 是3,而在直线,$n = 1$.用 $\mathbf{R}$ 表示全体实数,用 $\mathbf{R}^n$ 表示实数的全体n 组.我们将几何地看待这些;于是 $\mathbf{R} = \mathbf{R}^1$ 是数直线(有坐标系的直线),$\mathbf{R}^2$ 表示一个有坐标系的平面,而 $\mathbf{R}^3$ 表示一个有坐标系的三维空间.

从前面所提出的计划,已经可以知道我们的重点将是函数的研究.本文中"函数"这名词的含义,比在较初等的课程(包括微积分在内)中通常含义要广泛一些.在本节中,我们将指出这种广泛性的范围.本文将采用集合论中的通用语言及符号.我们并不采用许多术语,但是采用了的就经常用.现在就来给出将来所需要的术语与符号.

若 X 是一个集合,则 $x \in X$ 表示x 是这集合的一个元素.若 A 与X 都是集合,则 $A \subset X$(读作:A 被包含在X 中),表示 A 的每个元素是X 的一个元素,而 A 叫做X 的一个子集.大多数情况下,我们要涉及的将是直线的,或平面的,或空间的子集(即 $X \subset \mathbf{R}^n$,对于某个正整数 n),所以总是把 X 的元素叫做点.若 $A \subset X$ 而且$B \subset X$,则它们的交集 $A \cap B$ 由A 与B 的公共点的全体组成,而它们的并集 $A \cup B$ 由A 的点、B 的点以及A 与B 的公共点的全体组成.空集用$\varnothing$ 表示,它是任何集合的子集.于是 $A \cap B = \varnothing$ 是说 A 与B 没有公共点.若 $A \subset X$,则 $X - A$ 表示在X 中,但不在 A 中的点的全体,叫做 A 在X 中的补集.

在高等数学里,函数这名词的意义十分广泛;事实上,它是整个数学的一个基本概念.下面的定义,就与函数的这种最广泛的用法相一致.

定义 一个函数f由三件东西组成:一个集合 X,叫做f 的定义域;一个集合 Y,叫做f 的值域;与一个规律,它对于 X 的每一个元素,指定 Y 的一个相对应的元素.符号 $f: X \to Y$.读作:f 是一个函数,它的定义域为 X 而值域为 Y;或简单些,f是从X 到Y 的一个函数.我们把"f对于每一个元素$x \in X$,指定一个对应的元素 $y \in Y$"这句话缩写成"$y = fx, \forall x \in X$."(我们把 $f(x)$ 中括号省略掉,以符合通行的用法).对于 $x \in X$,全体 fx 值的集合正好是Y 时,我们说f 是

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

从 X 到 Y 上(onto)的一个函数,或从 X 到 Y 的一个满函数①.

例如在微积分中,一个函数通常是指实变数的实值函数;即它的定义域与值域都是 $\mathbf{R}$ 的子集.此外,还经常假设函数是由像$\sqrt{1-x}$这样的某一公式给出来的.这种情况下,照例即不说明定义域也不说明值域.默认定义域 X 是使公式有意义的那些实数的集合(例如$\sqrt{1-x}$ 对于所有的 $x \leqslant 1$ 包含负数,都有定义).作为值域 Y,总是取所有函数值的集合(例如,对于$\sqrt{1-x}$,它是所有的 $y \geqslant 0$ 这集合).这种函数叫做数值函数.

在更高深的书中,不能假设 f 由一公式给出,以及从公式能得出定义域与值域.此外,我们也不愿意把 X 与 Y 局限为 $\mathbf{R}$ 的子集.在本文中,X 与 Y 经常会是不同维的欧几里得空间的子集:$X \subset \mathbf{R}^m$ 而 $Y \subset \mathbf{R}^n$.所以我们必须小心区别每一个函数的 X 与 Y.我们也不总是假设 Y 恰好是函数值的集合;它可能大一些.让我们首先考虑某些较熟悉的、几何地定义的函数.

平面的一个平移是一函数 $f:\mathbf{R}^2 \to \mathbf{R}^2$.它是一种刚体的均匀运动的结果,其中每个点走过一线段或向量;向量是各点所走过的路径,一律平行,且有相同的长度与方向.只要用一点所描出的向量就能完全刻画一个平移,因为别个点的向量可以照样得出.于是,若已知 f 把点 p 带到点 q,则 f 将把点 p' 带到点 q' 使 p,q,p',q' 形成一个平行四边形.例如,若 f 把原点$(0,0)$带到$(2,-3)$,则它把(x_1,x_2)带到点(x_1+2,x_2-3).所以 f 的公式是:$y_1 = x_1 + 2$ 与 $y_2 = x_2 - 3$.

平面的一个旋转是一个函数 $f:\mathbf{R}^2 \to \mathbf{R}^2$,另一个刚体运动的结果;这次是围绕一个叫做旋转中心的定点 z 的转动.f 把每一个中心为 z 的圆周变成它自身,并且把每一条从 z 出发的射线(包含起点的半直线)变成另一条射线.这样的两射线间的夹角叫做旋转角,而且任一条射线都旋转同样的角度.中心与旋转角完全刻画了旋转.

$\mathbf{R}^2$ 的对于其中一条直线 L 的反射是一个函数 $f:\mathbf{R}^2 \to \mathbf{R}^2$;它是一个刚体运动,使 L 上的每个点固定,而使 L 的两侧互换.最形象的办法:把它看成是这平面在空间中绕直线 L 旋转 $180°$ 的结果.

平面变成它自身的任何一个刚体运动的结果是一个平移,或是一个旋转,或是一个反射,或是一个反射后又紧跟着一个平移;这个事实是能证明的.平面中图形的形状和大小不改变;只它们的位置与定向可能改变.这些函数都是初等几何中的全等.

---

① 我们把点 $y = fx$ 叫做点 x 的 f 象或象.很自然地,我们要用 fX 表示 X 中全体点 x 的象 fx 所形成的集合,叫做 X 的 f 象或象.这句话可以用集合论里的一个十分方便的符号:$fX = \{fx \mid \forall x \in X\}$ 来表示.这符号中花括号{}总是用来表示一个集合,括号内竖线的左边说明这集合的元素是 fx,右边说明这些元素是对于 X 的每一个元素 x 得到的.根据定义中对 f 的要求,下列事实是明显的.给定一点 x,恰只有一个象点 $y = fx \in Y$.在这意义下,f 也叫做单值函数.因此我们的定义只限于单值函数.而给定一点 $y \in Y$,并未要求存在点 $x \in X$ 或只有一点 $x \in X$,使得 $fx = y$.满足 $fx = y \in Y$ 的点 x 叫做 y 的原象或反象.
当 $Y = fx$,f 就叫做从 X 到 Y 的满函数.当 Y 的每一点 y,它的原象只有一个点,f 就叫做单原函数.单值的、单原的,而且满的函数就叫做从 X 到 Y 的一对一的函数,或 X 与 Y 间的一对一的函数.——译者注





平面的一个相似变换是一个函数 $f:\mathbf{R}^2\to\mathbf{R}^2$,它把所有长度都改变成 r 倍,这里 r 是一个常数.作为一个例子,取定 $\mathbf{R}^2$ 的一点 z,令 $fz=z$.并且,对于每一个其他的点 $x\neq z$,在从 z 到 x 的射线上,作从 z 开始的线段,使其长度是从 z 到 x 的线段(或向量)长度的二倍,然后定义 fx 是这个新作的线段(或向量)的终点.这个 f 所用以改变长度的倍数是 $r=2$.这样一个以 $r>1$ 为倍数的 f 叫做以 z 为中心的一个放大.当 $r<1$ 时,f 叫做一个收缩.$r=1$ 时的相似变换是上述刚体运动的一种.$r\neq 1$ 时的相似变换总有一个固定点 z,它是以 z 为中心的一个收缩或一个放大,再跟着作一个绕 z 的旋转或作一个对于过 z 的某条直线的反射.相似变换永远使直线变成直线,并且不改变线与线之间夹角的量度.它能改变一个图形的大小、位置与定向,却不改变它的形状.

令 L 为平面 $\mathbf{R}^2$ 中的一条直线,垂直射影 $f:\mathbf{R}^2\to L$,规定 $\mathbf{R}^2$ 的每一点 x 的象点 fx 是从 x 到 L 所引垂线的垂足.

令 S 为 $\mathbf{R}^3$ 中的以 z 为球心的一个球面.径向射影 $f:\mathbf{R}^3-z\to S$,规定 $\mathbf{R}^3$ 每一个不是 z 的点 x 的象点 fx 是从 z 到 x 的射线与 S 的交点.

前面的一些例子,部分地说明我们将感兴趣的那种函数.为了描绘这样的函数,并对它们进行有意义的叙述,我们采用象与反象或原象这两个概念.若 $f:X\to Y$ 与 $A\subset X$,则 A 在函数 f 作用下的象,或简单地说 A 的 f 象,用 fA 表示,它是 Y 的子集,由全体 $x\in A$ 的 fx 值组成.确切地说,一点 $y\in Y$ 是在 fA 内,指的是,至少有一个 $x\in A$,使 $fx=y$.人们能把 fA 看做是对于整个 A 运用 f 的结果.例如,在一刚体运动或一相似变换 $f:\mathbf{R}^2\to\mathbf{R}^2$ 的作用下,$\mathbf{R}^2$ 的任一直线 L 的象是 $\mathbf{R}^2$ 的一直线 fL.在垂直射影 $f:\mathbf{R}^2\to L$ 下,$\mathbf{R}^2$ 的每一线段 A 的象,是 L 上的一线段 fA(图 2.1),当 A 与 L 垂直时是 L 上的一个单独的点(图 2.2).在径向射影 $f:\mathbf{R}^3-z\to S$ 下,$\mathbf{R}^3$ 的不通过 z 的每一条直线 L 的象 fL 是 S 上一个大的半圆,不包含端点(图 2.3).

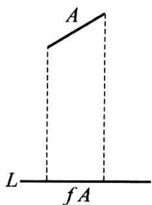


图 2.1

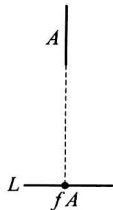


图 2.2

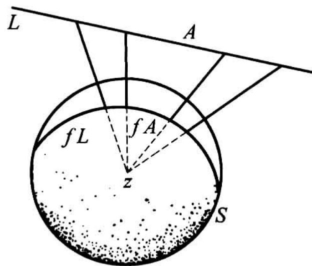


图 2.3

如果考虑 $\mathbf{R}^3$ 对于一个平面的任意反射,就明白我们用“象”这个字的来源.

如果 $f:X\to Y$ 与 $B\subset Y$,则使 $fx\in B$ 的那些点 x,形成 X 的一个子集,叫

做在 f 下的 B 的反象或原象，用 $f^{-1}B$ 表示. 在垂直射影 $\mathbf{R}^2 \to L$ 作用下，L 的单独一个点 y 的原象，是在 y 处与 L 垂直的直线；而线段的反象或原象，是该线段在其两端点处向 L 所引两垂线之间所夹的一长条. 在径向射影 $\mathbf{R}^3 - z \to S$ 下，S 的一点 y 的反象或原象是从 z 出发通过 y 而删去 z 的半直线. S 上一个圆域的反象或原象是挖去顶点的一个圆锥.

令 $f: X \to Y$ 且 $A \subset X$，则 A 的象也在 Y 中. 如果 Y 的一个子集 B 包含 fA，则对于所有的 $x \in A$，由 $gx = fx$ 定义的函数 $g: A \to B$ 叫做 f 在 A 与 B 上的限制. 最经常需要的是只限制 f 的定义域，在这种情况下，f 在 A 与 Y 上的限制用 $f \mid A$ 表示(读作：f 在 A 上的限制). 例如，如果 $f: \mathbf{R}^2 \to \mathbf{R}^2$ 是平移，而 L 是 $\mathbf{R}^2$ 中的一条直线，则 $f \mid L$ 是移 L 到与它平行的位置.

如果有两个函数 $f: X \to Y$ 与 $g: Y \to Z$，则能把这两个函数复合成一个新函数，用 $gf: X \to Z$ 表示；它对于 X 中的每个 x，在 Z 中规定 $g(fx)$ 这个元素. 例如，令 f 与 g 为平移 $\mathbf{R}^2 \to \mathbf{R}^2$，其中 f 把每个点向东移两个单位，而 g 把每个点向北移两个单位. 那么，gf 是把每个点向东北移动 $2\sqrt{2}$ 单位的平移. 作为第二个例子，令 f 与 g 为由公式

$$fx = x^2 + 1, gx = 2 - x$$

给出的数值函数 $\mathbf{R} \to \mathbf{R}$. 则能形成复合函数 gf 与 fg，由下列公式给出

$$gfx = g(fx) = 2 - (x^2 + 1) = 1 - x^2$$

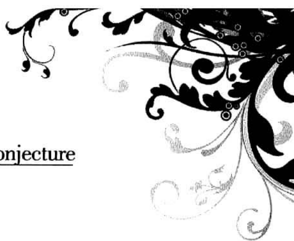
$$fgx = f(gx) = (2 - x)^2 + 1 = 5 - 4x + x^2$$

有些简单的函数如果不特别地提出来，人们会觉察不到它们的存在. 首先是常函数：一个函数 $f: X \to Y$ 是一个常函数，如果象 fX 是 Y 的一个单独的点. 对于 Y 的每一个点，有一个常函数. 其次，集 X 的恒同函数：对于每一个 $x \in X$，函数 $f: X \to X$ 都使 $fx = x$. 最后，若 $A \subset X$，对于每一个 $x \in A$，函数 $f: A \to X$ 使得 $fx = x$，则 f 叫做包含函数. 显然，包含函数是恒同函数在子集上的限制. 一个常函数的任一限制是常函数.

函数 $f: X \to Y$ 称为一对一的(简记为 1 - 1)，如果 Y 的每一点是 X 的一个点而且只是一个点在 f 下的象①，在这种情况时，如果对应于 Y 的每一点 y，指定 X 中的使 $fx = y$ 成立的唯一的点 x，就是从 Y 到 X 的一个函数，称为 f 的反函数，用符号 $f^{-1}: Y \to X$ 表示. 例如，平面的每一个刚体运动都是一对一的. 如果 f 是由以 p 与 q 为端点的线段规定的从 p 到 q 的平移，则 f^{-1} 是从 q 到 p 的平移. 旋转的反函数是具有同一个中心，而角的大小相同但方向相反的一个旋转. 平面的相似变换也是一对一的. 倍数为 $r > 1$ 的放大的反函数是具有同一中心而倍

① 有些作者称函数 $f: X \to Y$ 是一对一的，如果 Y 的每一点是 X 的至多一点的象，这样就允许 fX 可以不是 Y 的全体；而如果 $fX = Y$，则说 f 是一对一且满的.





数为$\frac{1}{r}$的收缩.

在处理数值函数时,人们企图从 $y = fx$ 求出f的反函数的公式,即用 y 表出的解 x. 于是,开平方根是平方的反函数,而对数是指数的反函数. 在某些点的原象不是单独的点的情况下,人们就把 f 限制在一个子集上,以获得一个一对一的函数. 例如,如果 $fx = x^2$,从 $y = x^2$ 得到 $x = \sqrt{y}$ 与 $x = -\sqrt{y}$,但如果把 f 的定义域限制在正数与零这子集 A 上,也把 f 的值域限制在同一个子集上,则这被限制的函数 g 是一对一的,而它的反函数是通常的方根函数 $g^{-1}y = \sqrt{y}$. 在指数函数 $fx = 10^n$ 的情形,只需要把 f 的值域限制在正数集上,即得一对一的函数.

虽然所列举的函数例子,已经各色各样,但它们仍不足以使人们理解到函数概念是怎样的广泛和基本. 作为广义的函数的一个例子,考虑"男孩的母亲"这概念. 男孩的集合为定义域 X,妇女的集合为值域 Y,对每个男孩 x 指定他的母亲为 fx. 我们的经历中充满着这样的例子: 书的颜色,屋子的顶,等. 什么时候用到"的"字,就出现一个函数. 名词的所有格出现时也如此.

科学中到处是函数. 化学反应的结果是放在一起的化学物的函数. 物理实验的结果是实验装置条件的函数.

回到数学方面来,处处都有"那个的这个"这样形式的话. 例如:圆的面积,线段的中点,角的平分线,两个集合的并集等. 这些例子的每一个都是一个函数. 在两集的并集情形下,函数定义域的一个元素,是已知集合 X 的一对子集(A,B),而它的值域是 X 的子集的集合.

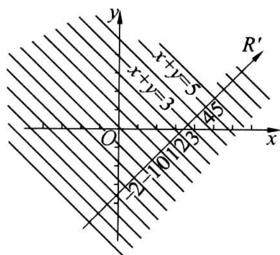


图 2.4

许多函数都能用几何图描绘出来. 例如,x 与 y 两数的和 $x + y$ 能看成是一个函数$f: \mathbf{R}^2 \to \mathbf{R}^1$,把每一对$(x,y)$看成是 $\mathbf{R}^2$ 中的一个点,f 就是平面到一条直线上的射影. 方程为 $x + y = 3$ 的直线是 $f^{-1}(3)$. 其他数的反象形成平行线族(图 2.4). 如果把值域 $\mathbf{R}^1$ 描绘成一条垂直于这平行族的直线,则 f 能看成为盖满这条直线的垂直射影.

3 邻域与连续

若 x 与 y 是 $\mathbf{R}^n$ 中的两个点,所谓它们之间的距离,指的是平常的直线距离,用 $d(x,y)$ 表示. 在用 x 与 y 的坐标时,能根据下列公式计算

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

这个公式是应用勾股定理的推广而获得的. 在 $n = 1$ 时,这公式简化为

$d(x,y) = |x-y|$（$x-y$ 的绝对值）. 现在我们不直接用这个公式，而只用距离函数 $d(x,y)$ 的能用这公式加以证明的某些性质. 这些性质都是众所周知的，现在把它们列举出来而不加证明.

首先，若 x 与 y 是不同的点，则 $d(x,y) > 0$. 其次，$d(x,x) = 0$ 对于所有的 x. 再次，对于每一对点 x 与 y，距离是对称的：$d(x,y) = d(y,x)$. 最后，对于任意三个点 x,y 与 z，有

$$d(x,z) \leqslant d(x,y) + d(y,z)$$

最后这个叫做“三角不等式”，因为它断言：三角形两边的长度之和大于或等于第三边的长度.

让我们回忆一下，在第 1 节中我们初步检验各种图形时，图形有否断裂对于我们的结论是否成立起着决定性作用. 就是因为这一原因，所以主要定理中的函数必须是连续的. 那时候我们用几何图形来直观地描述连续性；它的确切定义要用到另一个概念，即邻域的概念. 现在给出邻域的定义.

定义 令 X 为 $\mathbf{R}^n$ 的一个子集，x 为 X 中的一点，r 为一正实数. 然后，把 X 中的与 x 距离小于 r 的点的全体，定义为 x 的在 X 中的以 r 为半径的邻域，用 $N(x,r,X)$ 表示，而在 X 明白无误，不会引起误会时，可简单地用 $N(x,r)$ 表示.

例如，若 $X = \mathbf{R}^n$ 而 $n = 2$，则 $N(x,r,\mathbf{R}^2)$ 恰好是以 x 为中心，以 r 为半径的圆的内部. 同样，x 在 $\mathbf{R}^3$ 中的，半径为 r 的邻域是以 x 为中心，以 r 为半径的球的内部. 在 $n = 1$ 时，$N(x,r,\mathbf{R})$ 就是以 x 为中点，长度为 $2r$ 的开区间 $(x-r, x+r)$. 当 X 不是整个 $\mathbf{R}^n$ 时，则 $N(x,r,X)$ 恰就是 X 的位于 $N(x,r,\mathbf{R}^n)$ 内部的那部分（图 3.1），即 X 自身与 x 在 $\mathbf{R}^n$ 中的邻域 $N(x,r,\mathbf{R}^n)$ 的交集

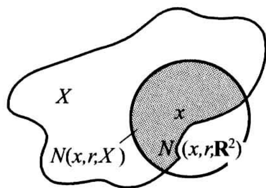


图 3.1

$$N(x,r,X) = X \cap N(x,r,\mathbf{R}^n)①$$

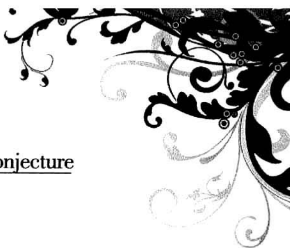
现在我们来谈函数的连续性这个重要概念.

定义 设 $f: X \to Y$ 是一个函数，其中 $X \subset \mathbf{R}^m$ 而 $Y \subset \mathbf{R}^n$，又设 $x \in X$. 如果对于 fx 的在 Y 中的每一个邻域，X 中存在 x 的某一个邻域，它的 f 象被包含在 fx 的所考虑的邻域中，则说 f 在 x 处连续.

为了简短地表达这个条件，按照微积分中惯用的记号，用 ε 与 δ 表示这些邻域的半径，把它们分别叫做 ε 邻域与 δ 邻域. 则这条件能重述如下：对于每一个

---

① 读者容易看出：当 $X = \mathbf{R}^n$ 时，这个式子也成立. 于是，不论 $X \subset \mathbf{R}^n$ 或 $X = \mathbf{R}^n$，对于 $x \in X$ 以及对于 x 的在 X 中的以 r 为半径的邻域 $N(x,r,X)$ 与在 $\mathbf{R}^n$ 中的以 r 为半径的邻域 $N(x,r,\mathbf{R}^n)$，这个式子都成立.——译者注



正数 ε,存在一个正数 δ,使 x 的 δ 邻域的 f 象,属于 fx 的 ε 邻域

$$fN(x,\delta,X) \subset N(fx,\varepsilon,Y)$$

若 f 在 X 的每一点处连续,则说函数 f 在 X 中连续.

若把 δ 和 ε 解释为接近的程度,则定义可以意译为:人们只需要求 x' 足够接近 x,就能使 fx' 接近 fx.更粗略的意译是:x 的微小改变,引起 fx 的微小改变.

拓扑学中只注意连续函数,并把"连续函数"这术语缩简为"映射",所以球极射影就是从挖去了极的球面到平面的映射.然而,为了说明连续性定义的意义,让我们来考察几个不连续函数的例子.

作为第一个例子,令 $f:\mathbf{R}^2 \to \mathbf{R}^2$ 使平面上除去一个点 p 外,每个点都固定不动,而令 $fp = q$ 为另一个点.为说得具体,可取 p 为坐标原点(0, 0)而 q 为(1,0).那么,f 在非 p 的每一点处连续.要看出它在 p 处不连续,取 ε 为 $d(p,q)$ 的一半,使得 $N(q,\varepsilon) = N(q,\frac{1}{2})$ 为以(1,0)为中心,以 $\frac{1}{2}$ 为半径的圆的内部(图 3.2).这时候,没有 p 的邻域会映到 $N(q,\frac{1}{2})$ 中,因为 p 的每个邻域都包含有不在 $N(q,\frac{1}{2})$ 中的点,而 f 又使这些点不动,从而这些点的 f 象不在 $N(q,\frac{1}{2})$ 中.(图3.2给出邻域 $N(p,\frac{1}{4})$,其中的点只有 p 的 f 象 $fp = q$ 在 $N(q,\frac{1}{2})$ 中)因为对于 $\varepsilon = \frac{1}{2}$,不存在相应的 $\delta > 0$ 使得 $fN(p,\delta) \subset N(q,\frac{1}{2})$,所以 f 在 p 处不连续.直观的几何图形是:f 把点 p 从平面挖出来,再把它粘到点 q 处.

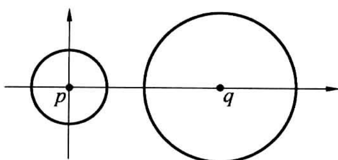


图 3.2

下面的第二个例子显示出不连续的另一种类型.把平面 $\mathbf{R}^2$ 分成两半(半平面)A 与 B,其中 A 包含所有使 $x_1 \geqslant 0$ 的点 (x_1,x_2),而 B 是它的补集.注意 $x_1 = 0$ 这条铅垂线 L 包含在 A 中(图 3.3).定义 f 为从 $\mathbf{R}^2$ 到 $\mathbf{R}^2$ 的这么一个函数,使得 f 在 A 上的限制是向右平移一个单位,而 $f|B$ 是向左平移一个单位.f 把两个

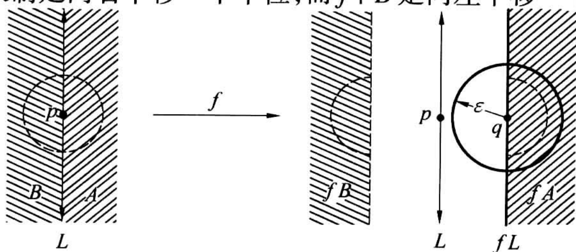


图 3.3

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

半平面 A 与 B 分裂开，A 移向右，B 移向左．因为 $L \subset A$，L 移向右．函数在 L 外的每点处都连续，而在 L 的每点处都不连续．为了证明后一断言，令 $p \in L$，令 $q = fp$，并令 ε 为小于 2 的任一正数．于是在 $\mathbf{R}^2$ 中 q 的以 ε 为半径的邻域，即 $N(q,\varepsilon)$，不包含 fB 的点，而每个 $N(p,r)$ 包点 B 的点．如图 3.3 所示，P 的任何邻域的左半（由虚点的圆表出）在 f 的作用下离开了 $N(q,\varepsilon)$．因而没有对应的 δ，使得 $\mathbf{R}^2$ 中的 p 的以 δ 为半径的邻域的象，包含在 q 的以 ε 为半径的邻域中，即使得

$$fN(p,\delta) \subset N(q,\varepsilon)$$

直观的图形是：f 把平面沿 L 割裂，再把两半分离开．

为着说明连续性的定义符合我们的直觉，还可用第 1 节中所举的不连续性的例子的图形来检验．在图 1.5 中，若邻域 $N(r,\varepsilon)$ 的半径 ε 小于距离 $d(r,s)$ 的一半，则 c 的任一邻域包含 X 的许多点，它们的象不在这个 $fc = r$ 的 ε 邻域中．同样，在图 1.6 中，若 ε 是比距离 $d(r,s)$ 与 $d(r,t)$ 中较小的还小，则 c 的任何邻域又包含着许多点，它们的象不在 $fc = r$ 的以 ε 为半径的邻域中．

要证明函数在某点不连续，只需要指出：对于单独的一个 $\varepsilon > 0$，不存在 δ．要证明连续，经常比较困难些，因为要指出：对于每个可能选出的 ε，怎样找出与它相对应的数 δ（就是必须显示出 δ 为 ε 的一个函数）．然而，对于一些简单函数，这并不困难；现在我们就来考虑这些函数．

对于任一个 $X \subset \mathbf{R}^n$，恒同函数 $f: X \to X$ 连续．回忆恒同函数的定义：对于所有的 $x \in X$，$fx = x$．对应于 $x \in X$ 与一个 $\varepsilon > 0$，取 $\delta = \varepsilon$．显然 f 把 $N(x,\delta,X)$ 映入 $N(fx,\varepsilon,X)$，因为 f 使每一点不动，故这两邻域重合．同样，若 $A \subset X$，则包含函数 $f: A \to X$ 连续；因为它是一个映射．还取 $\delta = \varepsilon$，并利用 $N(x,\delta,A) = A \cap N(x,\delta,X)$ 这一事实．

任意常函数 $f: X \to Y$ 连续．在这情形下，fX 为 Y 的一个单独的点，譬如说 q；因而对于每一个 x 与对于每一个 $\varepsilon > 0$，以及 $r > 0$，$fN(x,r,X) \subset N(q,\varepsilon,Y)$；故不妨取 $\delta = \varepsilon$．

任意一个刚体函数 $f: X \to Y$ 连续．所谓"刚体"函数，是指不改变距离的函数

$$d(fx,fx') = d(x,x') \quad 对于所有的\ x,x' \in X$$

例如，$\mathbf{R}^2$ 的平移、旋转与反射是刚体的．要证明在点 $x \in X$ 处连续，对于每个 $\varepsilon > 0$，可取 $\delta = \varepsilon$．然后，若 $x' \in N(x,\delta,X)$，则有 $d(x,x') < \varepsilon$，所以 $d(fx,fx') < \varepsilon$，因而 $fx' \in N(fx,\varepsilon,Y)$．即，因 f 的刚体性，x 的 ε 邻域的 f 象在 fx 的 ε 邻域中．

使距离紧缩的函数 $f: X \to Y$ 连续．紧缩这一条件由下式表出

$$d(fx,fx') \leqslant d(x,x') \quad 对于所有的\ x,x' \in X$$



再对于所有的 x,取 $\delta = \varepsilon$,并且应用上一段的论证.

任一相似函数 $f: X \to Y$ 连续.这里的条件是把所有的距离改变同一个倍数,譬如说,是 k 倍

$$d(fx, fx') = kd(x, x') \quad 对于所有的\ x, x' \in X$$

若 $0 \leqslant k \leqslant 1$,则 f 为压缩而可引用前一段.当 $k > 1$ 时,对于所有的点 x,取 $\delta = \frac{\varepsilon}{k}$.然后,$x' \in N(x, \frac{\varepsilon}{k}, X)$ 蕴涵 $d(x, x') < \frac{\varepsilon}{k}$.这能写为 $kd(x, x') < \varepsilon$.上面的等式就产生 $d(fx, fx') < \varepsilon$,从而 $fx' \in N(fx, \varepsilon, Y)$.例如,若 $k = 2$ 而 f 使距离双倍,则以邻域 $N(fx, \varepsilon, Y)$ 的半径之半为半径的 x 邻域的 f 象 $\subset$ 邻域 $N(fx, \varepsilon, Y)$.

所谓从球心 z 到球 S 的径向射影是函数 f: $\mathbf{R}^3 - z \to S$. f 是满的,是连续的. f 把 S 外部的两个点投射到 S 上,缩小了它们之间的距离,所以那限制在 S 外部的 f,显然连续.然而,限制在 S 内部的 f 放大距离;如果两点在球内移动,当它们离球心越近时,f 把它们之间的距离放大的倍数就越大.要证明在一点 $x \neq z$ 处连续,作为 ε 的函数的 δ 的表达式相当复杂,但注意到各种相似三角形以及用一点儿代数,仍然能得到这个表达式.我们也能几何地得出一个 δ 如下.令 t 为 S 与以 fx 为中心,以 ε 为半径的球的一个交点.图3.4显示出通过 x, fx 与 t 这三个点的平面的截面.令 δ 为从 x 到 zt 线的垂直距离.于是球 $N(x, \delta)$ 内部的每个点的投影落在 $N(fx, \varepsilon, S)$ 中.

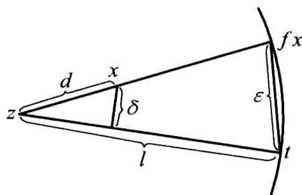


图 3.4

4 开集和闭集

设 X 是 $\mathbf{R}^n$ 的一个子集.我们的目的是要定义并研究 X 的一类特别的子集,即所谓的 X 的开集.它们将在随后的讨论中扮演主要角色,因为借助于它们,才能容易表达出所要讨论的 X 的各种拓扑性质.还有,用开集时,函数连续的条件取一个很简单的形式.

不容易事先看出:为什么开集这个概念会是一个重要概念.它逐渐得到承认,是一个历史事实.在拓扑发展的早期(1900 ~ 1930),对这一课题人们曾采取各种途径,做过各种不同的探讨.人们曾采用过诸如下列名词所表达的概念:邻域空间,度量空间,极限点,序列极限以及闭包.当时并未弄清楚这些途径都等价,都殊途同归;也无人能预言拓扑的发展方向与最终形式.直到这个时期的末期,才逐渐弄清楚下述事实:开集概念是探讨所有拓扑性质的简单而灵活的工

具.从那以后,人们公认开集这个概念所提供的途径最优越.

定义 令 X 为 $\mathbf{R}^n$ 的一个子集.若对于 X 的子集 U 的每一点 x, U 包含 x 的在 X 中的某邻域,则称子集 U 为 X 的开集.条件可重述如下:对于每一点 $x \in U$,有一个数 $r > 0$,使 $N(x, r, X) \subset U$.

我们不久会看到,开集是易于遇到的,并且以各式各样的形式出现.我们的第一类例子,就是 X 中的点的邻域:凡邻域都是开集.

设已知 $x_0 \in X$ 与 $r_0 > 0$.要证明上述断言,必须证明 $N(x_0, r_0, X)$ 是 X 的开集(图 4.1);这就需要指出;对于每一点 $x \in N(x_0, r_0, X)$,存在一正数 r,使邻域 $N(x, r, X) \subset N(x_0, r_0, X)$. 令 $r = r_0 - d(x_1, x_0)$. 现在 $x \in N(x_0, r_0, X)$ 指的是 $d(x, x_0) < r_0$;因此 r 必为正数.令 $y \in N(x, r, X)$,于是 $d(x, y) < r$.从三角不等式可知

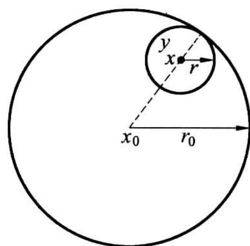


图 4.1

$$d(x_0, y) \leqslant d(x_0, x) + d(x, y)$$

又因 $d(x, y) < r$,有

$$d(x_0, y) < d(x_0, x) + r$$

但由我们的 r 的定义,$d(x_0, x) + r = r_0$,因此

$$d(x_0, y) < r_0$$

这就表明 $N(x, r, X)$ 的任一点 y 是 $N(x_0, r_0, X)$ 的一个点,所以 $N(x_0, r_0, X)$ 是 X 的开集.

下面的一些定理说明怎样从已知的开集作出另外的开集.

定理4.1 若 U 与 V 都是 X 的开集,则它们的交集 $U \cap V$ 是 X 的一个开集. X 的任何有限个开集的交集是 X 的一个开集.

为着证明第一个断言,令 $x \in U \cap V$.因 $x \in U$ 且 U 是开集,有一个 $r > 0$ 使 $N(x, r, X) \subset U$.因 $x \in V$ 而 V 是开集,有一个 $s > 0$ 使 $N(x, s, X) \subset V$.设 t 为 r 与 s 中较小的那个;显然 $N(x, t, X)$ 在 U 与 V 中,故在 $U \cap V$ 中.这就证明 $U \cap V$ 是开集.为着证明第二个断言,假设 x 在每一个开集 $U_1, U_2, \cdots, U_k$ 中.于是有数 $r_i > 0 (i = 1, 2, \cdots, k)$ 使 $N(x, r_i, X) \subset U_i$.取这些 $r_1, r_2, \cdots, r_k$ 中最小的一个,设为 t.显然 $N(x, t, X)$ 在交集中,这就完成了证明.

在上面的论证中,我们假设 $U \cap V$ 中有点 x.若 U 与 V 无公共点,即

$$U \cap V = \varnothing = \text{空集}$$

我们还必须根据开集的定义来验证空集是开集.要求这样的验证,乍一看,似乎有些迂腐气;但却是严格地符合逻辑的.因 $\varnothing$ 没有点,所以它的每个点有一个邻域包含在 $\varnothing$ 中.这说法是正确的,所以 $\varnothing$ 是开集.从相反的角度想一想.若集合 A 不是 X 中的开集,A 就包含某个点,而不包含该点(在 X 中)的邻域;所以一



个非开集是一个非空集.这一事实极其重要,它连同另一个重要事实,即一集合是它自己的开集,足够正式作为一个定理.

定理 4.2 空集$\varnothing$是 X 的一个开集,并且 X 自身是X 的一个开集.

第二个命题显而易见.因由邻域的定义,对于每个 $x \in X$,和所有的 $r > 0$

$$N(x,r,X) \subset X$$

下面的定理给出从已知开集建立新开集的另一个方法.

定理 4.3 X 的任意多个(有限个或无穷多个)开集的并集是 X 的一个开集.

为了证明,令 C 表示这么多个开集的全体,又令 A 表示它们的并集.若 $x \in A$,则 C 中存在一个开集U包含x.因 U 是开集,有一个 $r > 0$使 $N(x,r,X) \subset U$.由并集的定义,$U \subset A$,从而 $N(x,r,X) \subset A$,这就证明了 A 是开集.

这些结果表示,对于大多数集合 X,它的开集族是很庞大的.通过作邻域的并集能无止境地作成各式各样的开集.现在我们将说明许多熟知的集合,具有开集的资格.

若 $X = \mathbf{R}$,则每个邻域是一个开区间

$$N(x,r) = (x - r, x + r)$$

每个开区间是它的中点的一个邻域,因而是 $\mathbf{R}$ 的一个开集.两个或更多的开区间的并集,也是 $\mathbf{R}$ 的一个开集;例如,不相交的开区间的序列$\left(\frac{1}{n+1}, \frac{1}{n}\right)$,$n = 1,2,3,\cdots$,的并集是开集.

令 $X = \mathbf{R}^2$.这时候,凡圆的内部都是邻域,所以每一个圆的内部是$\mathbf{R}^2$ 中的开集.令 A 表示$\mathbf{R}^2$ 中的一个矩形的四条边(四条线段)上诸点的集合.它的补集 $\mathbf{R}^2 - A$ 有两部分:内部 U 与外部V(图4.2),若 x 是U的一个点,并选取一个正数 r,使它比从 x 到A 的边的最短距离还小,则 $N(x,r)$ 在 U 内.所以 U 是 $\mathbf{R}^2$ 中的开集.同样,外部 V是 $\mathbf{R}^2$ 中的开集;但 A 在 $\mathbf{R}^2$ 中不是开集;因为它的一个点 z,没有邻域 $N(z,r,\mathbf{R}^2) \subset A$.实际上,$A$ 的每个点都有这个性质.若用任意简单闭多边形如三角形或六边形来代替矩形 A,这些结论照样有效.

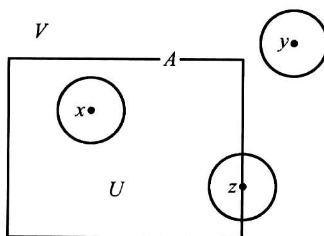


图 4.2

应该注意:开集这性质是脆弱的,即一个开集可以在添加一个点后,不再是开集.在上面的例中,若把 A 的内部U添加上A 的或A 的外部的一个点,则这扩大的集合,就不再是 $\mathbf{R}^2$ 的开集了.

当 $X = \mathbf{R}^3$ 时,邻域都是球的内部,而每一个这样的集合是在 $\mathbf{R}^3$ 中的开集.球的外部同样是 $\mathbf{R}^3$ 中的开集,但球面在 $\mathbf{R}^3$ 中就不是开集.令 A 表示 $\mathbf{R}^3$ 中一个矩形盒子的面、边与顶点的全体点,则 $\mathbf{R}^3 - A$ 分成两个开集:盒的内部与外部.令 T 表示 $\mathbf{R}^3$ 中一个环面(油炸圈饼)的面上的全体点,于是 $\mathbf{R}^3 - T$ 分成两个开集:T 的内部与外部.

在上面所举的例中,我们都取 $X = \mathbf{R}^n$.下面的定理告诉我们,一旦 $\mathbf{R}^n$ 的开集描绘清楚了,如何"看出"$\mathbf{R}^n$ 的一个子集 X 的开集.

定理 4.4 若 $X \subset \mathbf{R}^n$,则 X 的开集族就是 X 自身交 $\mathbf{R}^n$ 的全体开集的交集族①.

先证明:若 U 是 $\mathbf{R}^n$ 的一个开集,则 $X \cap U$ 是 X 的一个开集.令 $x \in X \cap U$.因 $x \in U$ 而 U 是 $\mathbf{R}^n$ 的开集,就有一个 $r > 0$ 使 $N(x,r) \subset U$.因而

$$X \cap N(x,r) \subset X \cap U$$

但 $N(x,r,X) = X \cap N(x,r)$,故

$$N(x,r,X) \subset X \cap U$$

这就是说:对于每个 $x \in X \cap U$,X 中有 x 的一个邻域属于 $X \cap U$,所以 $X \cap U$ 是 X 的一个开集.这就完成了定理的证明的一半.

其次,必须证明能把 X 的一个开集 V 放大成 $\mathbf{R}^n$ 的一个开集 U,使 $V = X \cap U$.如果 V 原来就是一个邻域 $N(x,r,X)$,所寻找的放大体 U 显然就是 $N(x,r,\mathbf{R}^n)$.现在,容易验证 X 的任一开集 V 是包含在 V 中的所有邻域的并集,所以我们只要寻找出那包含在 V 中的每个(形如 $N(x,r,X)$ 的)邻域的放大体($N(x,r,\mathbf{R}^n)$),来构造所寻找的 V 的放大体.对于所有使得 $N(x,r,X) \subset V$ 的 $x \in V$ 和 $r > 0$,我们定义 U 为邻域 $N(x,r,\mathbf{R}^n)$ 全体的并集,并且证明这个 U 满足要求,即证明 U 是 $\mathbf{R}^n$ 的开集,与 $X \cap U = V$.因为每一个 $N(x,r,\mathbf{R}^n)$ 是 $\mathbf{R}^n$ 的开集,定理 4.3 断言 U 是 $\mathbf{R}^n$ 的开集.证明 $X \cap U = V$ 还要用两个步骤来完成:首先要证明 V 的每一个元素是 $X \cap U$ 的一个元素,然后要证明 $X \cap U$ 的每一个元素是 V 的一个元素.

因为 V 是 X 的开集,每一点 $x \in V$ 有一 $N(x,r,X) \subset V$,所以 $x \in U$.这证明了 $V \subset U$.但 V 也是 X 的一个子集,所以 $V \subset X \cap U$.这完成了第一步.第二步要证明 $X \cap U \subset V$.先注意到任一点 $y \in X \cap U$ 既在 X 中又在 U 中.作为 U 的一个点,y 属于某个 $N(x,r,\mathbf{R}^n)$ 使 $N(x,r,X) \subset V$.因为 y 也在 X 中,它就在 $X \cap N(x,r,\mathbf{R}^n) = N(x,r,X)$ 中,又因为 $N(x,r,X) \subset V$,从而 $y \in V$.这就

① 请注意,定理 4.4 实际是说:若 $X \subset \mathbf{R}^n$,则 X 自身与 $\mathbf{R}^n$ 的任一开集的交集是 X 的一个开集;而且反之,X 的每一开集是 X 自身与 $\mathbf{R}^n$ 的某一开集的交集.—— 译者注





完成了 $X \cap U = V$ 的证明,并且也就完成了定理的证明.

让我们总起来看一下本节中所已做到的事. X 的一个开集 U 的主要性质是给它下定义的性质:每一个 $x \in U$ 有一个邻域 $N(x,r,X) \subset U$. 关于单独的一个开集,所能说的只是这一句话,并无其他. 我们的定理叙述了 X 的开集全族(即 X 的全体开集)的性质:开集全族的元素包括空集、X 自身与每个邻域 $N(x,r,X)$;此外,还包括它的任意有限个元素的交集,以及任意有限个或无穷多个元素的并集.

现在转到闭集的概念.

定义 令 X 为 $\mathbf{R}^n$ 的一个子集. X 的子集 A 叫做 X 的一个闭集,如果 A 在 X 中的补集 $X - A$ 是 X 的一个开集. 简短地说,A 是闭集如果 $X - A$ 是开集.

引用开集的定义,就得出 A 是 X 的闭集的下列检验法:$X - A$ 的每一点有一个不与 A 相交的邻域. 例如,由单独一个点所组成的集合,总是任一稍大集合 X 中的闭集;因为,若 x 是这个单独的点,而 y 是任一别的点,若 r 小于或等于从 x 到 y 的距离,则 $N(y,r)$ 不包含 x. 同样地,一平面或空间中的一直线 L,是任一较大集合中的一个闭集;因为,若 y 不在 L 上,而 r 是从 y 到 L 的最近点的距离,则 $N(y,r)$ 与 L 不相交.

从所举的每个开集,作它的补集,就得到一个闭集的例子. 在图 4.2 矩形 A 的例中,外部 V 的补集是 A 与其内部 U 的并集. 因 V 是一个开集,从而 $A \cup U$ 是 $\mathbf{R}^2$ 的一个闭集. 同样,$A \cup V$ 是 $\mathbf{R}^2$ 的一个闭集. 因 $U \cup V$ 是开集而 A 是 $U \cup V$ 的补集,从而 A 是 $\mathbf{R}^2$ 的闭集.

X 中的一个集合 A 与它的补集 $X - A$ 之间的关系是互逆的:$X - A$ 的补集是 A. 在 X 的子集之间的这种对应关系叫做 X 中的对偶性. 开集和闭集是对偶概念,因为一个开集的对偶是一个闭集,并且反过来也对.

由于开集与闭集之间的这种对偶性,我们能从已证明的关于开集的每个定理推演出关于闭集的一个正确的"对偶"定理. 在作出对偶命题时,重要的办法是利用下述事实:并集和交集在下述意义下是"对偶"运算. A 与 B 的并集的补集,是 A 的与 B 的补集的交集

$$X - (A \cup B) = (X - A) \cap (X - B)$$

同样地,A 与 B 的交集的补集,是 A 的与 B 的补集的并集

$$X - (A \cap B) = (X - A) \cup (X - B)$$

于是下面四个关于闭集的定理,对应于那些关于开集已证明的定理. 关于两个开集的交集定理,给出对偶

定理 4.1′ 若 A 与 B 都是 X 的闭集,则它们的并集 $A \cup B$ 是 X 的一个闭集. X 的任意有限个闭集的并集是 X 的一个闭集.

要获得 $\varnothing$ 与 X 都是开集的对偶命题,只需要注意到 $\varnothing$ 与 X 在 X 中互为补

集：$X-\varnothing=X$，而 $X-X=\varnothing$.

定理4.2′ 空集$\varnothing$与 X 自身这两个集合，既都是 X 的开集，又都是 X 的闭集.

X 的一个开集通常不是 X 的一个闭集，反过来也一样. 在第7节中将对于 X 中的既开且闭的子集进行仔细研究.

关于任意多个开集的并集的定理，给出它的对偶

定理4.3′ X 的任意多个（有限多个或无穷多个）闭集的交集，是 X 的一个闭集.

定理4.4（X 的开集族是 X 自身与 $\mathbf{R}^n$ 的开集族的交集族）的对偶命题是

定理4.4′ 若 $X\subset\mathbf{R}^n$，则 X 的闭集族就是 X 自身与 $\mathbf{R}^n$ 的全体闭集的交集族.

设 $A\subset X\subset\mathbf{R}^n$，而 A 是 $\mathbf{R}^n$ 的闭集，则本定理断言 $A\cap X$ 是 X 的闭集. 因 $A\cap X=A$，得

推论 设 A 是 $\mathbf{R}^n$ 的一个闭集. 如果 $\mathbf{R}^n$ 的子集 X 包含 A，则 A 是 X 的一个闭集.

不要认为 $\mathbf{R}^n$ 的每一个子集或者是 $\mathbf{R}^n$ 的开集，或者是 $\mathbf{R}^n$ 的闭集；许多集合两者都不是. 半开区间 $(a,b]$ 在 $\mathbf{R}$ 中既不是开集又不是闭集. $(a,b]$ 的异于 b 的任一点 x，就有某一邻域属于这区间. 另一方面，b 的每一个邻域包含着 $(a,b]$ 外的点. 在图4.2的例中，内部 U 加上 A 的单独一个点，既不是 $\mathbf{R}^2$ 的开集也不是它的闭集，前面提到过它不是开集. 它也不是闭集，因为 A 的任一点 x 的每一个邻域 $N(x,r)$ 包含 U 的点. 但是由检验一个集合为闭集的标准，它的补集的每一点必有一邻域不与它相交.

现在转到用开集概念来阐述函数的连续性. 阐述的简易，已足以表明将来研究连续问题时，开集多么有用.

定理4.5 函数 $f:X\to Y$ 连续的充分必要条件是 Y 的每个开集的原象是 X 的一个开集. 等价地，f 连续的充分必要条件是 Y 的每个闭集的原象是 X 的一个闭集.

回忆第3节中函数 f 连续的条件：对于每个 $x\in X$ 与每个 $\varepsilon>0$，存在一个 $\delta>0$，使 x 的 δ 邻域的 f 象属于 fx 的 ε 邻域. 现在，定理中的条件是：对于 Y 的每个开集 V，$f^{-1}V$ 是 X 的开集. 这一条件的叙述无疑地比较简单.

要证明这定理，先设 f 连续，并且 V 是 Y 的一个开集. 我们必须证明：每个点 $x\in f^{-1}V$ 有一个邻域属于 $f^{-1}V$. 现在 $fx\in V$，而且 V 是 Y 的开集，所以有一个数 $\varepsilon>0$ 使 $N(fx,\varepsilon,Y)\subset V$. 因 f 连续，存在一个数 $\delta>0$，使 $N(x,\delta,X)$ 的 f 象在 $N(fx,\varepsilon,Y)$ 中，从而在 V 中. 所以 $N(x,\delta,X)\subset f^{-1}V$，而这就证明了对于 Y 的每个开集 V，$f^{-1}V$ 是开集.





为证充分性,设f具有性质:对于Y的每一个开集V,$f^{-1}V$是开集;然后证明f连续.令$x\in X$,并令$\varepsilon>0$,$N(fx,\varepsilon,Y)$是Y的一个开集,所以它的在X中的原象$f^{-1}N(fx,\varepsilon,Y)$是$X$的开集;记$U=f^{-1}N(fx,\varepsilon,Y)$.因$x\in U$并且$U$是开集,$x$有某邻域$N(x,\delta,X)\subset U$.从而$fN(x,\delta,X)\subset N(fx,\varepsilon,Y)$,这就证明了$f$连续.

已经证明定理中关于开集的部分.关于闭集的对偶部分是下列事实的结果:对于任意函数:$f:X\to Y$,Y的一个子集A的原象在X中的补集即A在Y中的补集的原象.用符号表示,对于每个子集$A\subset Y$

$$X-f^{-1}A=f^{-1}(Y-A)$$

这公式的证明作为简短的练习,留给读者.让我们承认这个公式,并设f连续以及A是Y的闭集.从而$Y-A$是Y的开集.由定理的第一部分,$f^{-1}(Y-A)$是X的开集.它在X中的补集因而是X的闭集.上述公式告诉我们,这个补集是$f^{-1}A$;因而$f^{-1}A$是X的闭集.

为着完成本定理的证明,再设每个闭集的原象是一个闭集.设A是Y的一个开集,则$Y-A$是Y的闭集;因此$f^{-1}(Y-A)$是X的闭集.所以它的在X中的补集是X的开集.上述公式说这补集就是$f^{-1}A$.于是,对于每个开集A,$f^{-1}A$是开集.所以f连续.这就完成了证明.

定理 4.6 若$f:X\to Y$与$g:Y\to Z$都是连续函数,则它们的复合函数$gf:X\to Z$也连续.

令W为Z的一个开集.因g连续,前一个定理断言:$g^{-1}W$是Y的一个开集.又因f连续,它断言:$f^{-1}(g^{-1}W)$是X的一个开集.留给读者来证明

$$(gf)^{-1}W=f^{-1}(g^{-1}W)$$

作为简短的练习.于是,对于Z的每个开集W证明了$(gf)^{-1}W$是开集.根据前面的定理,这就说明gf连续.

在后继的各节中经常用"一个空间X"的说法.虽然毫无例外地,X还是某个$\mathbf{R}^n$的子集;但是,我们的观点是,只提X的点与X的开集,不问周围的空间$\mathbf{R}^n$.这叫做内蕴的观点.请读者复习本节中的定义与定理,并且注意除定理4.4与4.4′外,所用的都是内蕴言词.在第8节中将讨论这个观点的重要性.

5 实数系的完全性

本节的要点是论述:存在着足够多的实数.更仔细地说,如果把实数看做是能用十进制小数(位数有限或无穷多)表示出来的数,则实数已足够完全填满我们的数直线.

数学史记载着数系不断发展的过程.首先,史前的人类用一、二、三与多来

计数.接着是正整数的无穷序列的概念,连同它们的名称与缩写的记号.其次出现分数或有理数,接着是代数方程的"根"或代数数,然后是零与负数,最后是超越数.①

在各个阶段中,某些使用到数的人,逐渐意识到他们所了解的数的概念有缺陷,已不够用.几经尝试以后,他们终于成功地创造了新的数;把新数与旧数合在一起,就弥补了那种缺陷.绝大多数人都充分地了解我们需要整数与有理数以及它们的负数与零.很好地了解到为什么这些还不够的人就少些.

毕达哥拉斯学派首先发现$\sqrt{2}$不是一个有理数;明确地说,即没有一个分数,它的平方是2.这里是欧几里得的证明.用反证法,设$\frac{m}{n}$是一个分数,它的平方为2.实际能假设$\frac{m}{n}$是既约分数,就是说:m与n没有不等于1的整数的公因数.特别是,它们都不会同是偶数(所有的公因数2等都已被"消去").把等式$(\frac{m}{n})^2=2$写成$m^2=2n^2$,这就说明m^2是一个偶整数.现在,一个奇整数的平方仍是奇数

$$(2r+1)^2=4r^2+4r+1=2(2r^2+2r)+1$$

因m^2是偶数,从而m是偶数,所以$m=2k$,其中k是某个整数.把m的这个值代入$m^2=2n^2$中,变成$4k^2=2n^2$,即$n^2=2k^2$.这表示n^2是偶数,因而n是偶数.于是,m与n都是偶数,与$\frac{m}{n}$为既约分数的事实矛盾.这矛盾表明了不存在平方是2的分数.

因为毕达哥拉斯学派的人都是几何学者,他们需要$\sqrt{2}$这个数.他们从长度为d的一线段出发,利用直尺与圆规,可以作一个边长为d的正方形.由勾股定理,对角线的长应为$\sqrt{2}d$.

用下面的方法来想象数系的逐步发展.用一条直线L与L上叫做0与1的两个点开始,人们能利用圆规挨个儿标出其余的整数点$2,3,\cdots$与$-1,-2,\cdots$.利用另一个涉及作辅助线的作图法,能把每个区间$[n,n+1]$按所要求的分成许多等份.于是,从0与1两点开始,L上的坐标为有理数的点,都能作出来.

现在,在L上稠密地散布着有理点(稠密的意义是指:两个这样的点之间还均匀地散布着无穷多这样的点).容易理解人们会从而不知不觉地设想这些有理点就是L上的所有的点.但是,毕达哥拉斯学派的人发现,边长为1的正方形的对角线,将它在L上标出时(图5.1)所得的点$\sqrt{2}$,却不是这些有理点中的一个.这个发现对于发现者该是怎样一个震动啊!这促使人们创造新的数,与由这种几何作图所产生的L上的新点相对应.

① 关于数的概念的发展的详细论述,见《新数学丛书》中的第一卷,Ivan Niven的:Numbers:Rational and Irrational.



不幸,有理数平方根的发现,仍没有给出足够的数.三等分一个角或两倍一个立方体需要取有理数的立方根,而这些经常不是有理数的平方根.所以数学家被迫去创造有理数的 n 次根以及甚至于更多的代数数.代数数是整系数多项式方程的根.

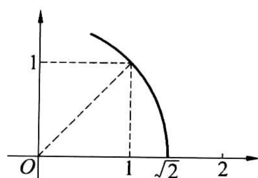


图 5.1

约在一百年以前人们发现代数数还不够;就是说,数直线上还有点不跟代数数相对应.特别是,曾经证明 π 这个数(一个圆的圆周与它的直径之比)不是一个代数数,人们不禁要问:这个发展过程有没有终止的时候;如果有,什么时候结束?

十进制与数的十进小数展开的发展,赋予这些问题一个新的观点.就已被创造出的所有的数而论,它们都能用它们的小数展开式来表示.应该着重指出,有理数的绝大多数以及所有其他的数都有无终止的小数展开式.在这阶段,自然地会想到把事情倒过来看;采取下述看法:任一小数展开式代表一个实数.那就是说,可以把实数集合 **R** 定义为小数展开式的集合.(依惯例约定:把后面都是九的小数展开式与后面都是零的展开式看做是同一个数,例如 $3.269\,99\cdots = 3.270\,00\cdots$)这实际正是我们所要做的.要说明这定义正确,我们必须表明它终止了创造新数;集合 **R** 的数完全填满了数直线.

必须解释"填满这直线"的意义.回忆求平方根例如$\sqrt{2}$的标准方法.从几何图形看,这是用方程 $y = x^2$ 的图象,并试着去确定它与水平线 $y = 2$ 的交点的横坐标(图 5.2).首先检验头几个整数,看出 $1^2 = 1$ 太小,而 $2^2 = 4$ 又太大.若 x 从一个正数递增到另一个,它的平方也随着增加.这个事实告诉我们,$\sqrt{2}$ 在区间 $I_0 = [1,2]$ 的某一处,而它的小数展开式的整数部分是 1.其次,把区间 I_0 分成十份,并计算 1.0,1.1,1.2,…,1.9,2.0 各数的平方.找到$(1.4)^2$小于 2 而$(1.5)^2$大于 2.于是$\sqrt{2}$在区间 $I_1 = [1.4,1.5]$ 中,而$\sqrt{2}$的小数展开式的开头是 1.4.再把 I_1 分成十等份,检验每个分点的平方数而找到$\sqrt{2}$ 在区间 $I_2 = [1.41,1.42]$ 之内.这样继续下去,就确定出一个无穷区间序列 $I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_k \supset \cdots$,挤近在$\sqrt{2}$处.($\supset$ 是 $\subset$ 的反向,而"$A \supset B$"读作"A 包含 B")每一个区间是它前一个的十分之一,而$\sqrt{2}$ 的小数展开式能从区间的左端点的小数展开式读出来,先是 1,然后是 1.4,再后是 1,41,等.

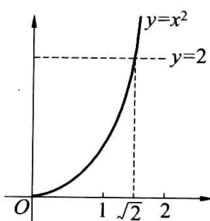


图 5.2

设想一个类似的,但更普遍的问题.不限于考虑 $y = x^2$ 而考虑 $y = fx$,其中 f 是任一连续函数,它随 x 的递增而递增;并且不限于找一数 x 使$x^2 = 2$,而求解方程 $fx = b$,其中 b 是某个已知数.若能找出一个起始的区间 $I_0 = [n, n+1]$,使 $fn < b$ 而 $f(n+1) > b$,那就能进行一再十等分区间的步骤.这给出一个无

穷的区间序列 $I_0 \supset I_1 \supset I_2 \supset \cdots \supset I_k \supset \cdots$. 正如找$\sqrt{2}$这个数的情形一样,综合这些区间左端点的小数展开式,能得出一个数 a 的小数展开式,数 a 在每一个这种区间中,所以应该是我们的问题:$fa = b$ 的一个解. 这就强有力地提示出如下的结论:如果我们同意每个小数展开式确定一个实数,那就有足够多的实数 a 来解出上述类型的任何问题.

现在我们来提出与证明一个定理,确切地重述前面的思想.

定义 实数闭区间的一个无穷序列 $I_0, I_1, I_2, I_3, \cdots, I_n, \cdots$ 叫做一个收缩序列,如果每个区间包含下一个区间,因而包含所有跟在后面的区间

$$I_0 \supset I_1 \supset I_2 \supset I_3 \supset \cdots \supset I_n \supset \cdots$$

一个收缩序列叫做一个正则收缩序列,若 $I_0 = [m, m+1]$ 是一个整数 m 到下一个整数 $m+1$ 的闭区间,并且,对于每个 $n \geqslant 1$, I_n 是 I_{n-1} 的十等份中的一个区间. 于是 I_1 是 I_0 的十分之一,I_2 是 I_1 的十分之一,照此类推. I_n 的长度是 10^{-n}.

完全性定理 每一个收缩区间序列有一个公共点;就是说,这一族所有区间的交集是非空的.

首先考虑正则收缩序列的情形. 对于每个 $n = 0, 1, 2, \cdots$,令 a_n 表示 I_n 的左端点;于是 $a_0 = m$ 是一个整数. 因 I_n 的长度是 10^{-n}, $I_n = [a_n, a_n + 10^{-n}]$. 把 I_{n-1} 分成十等份的点是

$$a_{n-1}, a_{n-1} + \frac{1}{10^n}, a_{n-1} + \frac{2}{10^n}, \cdots, a_{n-1} + \frac{9}{10^n}, a_{n-1} + \frac{10}{10^n}$$

因 I_n 是这十个子区间中的一个,它的左端点应该是

$$a_n = a_{n-1} + \frac{k_n}{10^n}$$

其中 k_n 是数字 $0, 1, 2, \cdots, 9$ 中的一个. 按照这种方式,这区间序列决定一个整数 m 与数字的一个无穷序列 $k_1, k_2, \cdots, k_n, \cdots$. 令 c 表示实数

$$c = m + \frac{k_1}{10} + \frac{k_2}{10^2} + \cdots + \frac{k_n}{10^n} + \cdots$$

(即 c 的小数展开式是 $m.k_1k_2k_3\cdots$). 因

$$a_n = m + \frac{k_1}{10} + \frac{k_2}{10^2} + \cdots + \frac{k_n}{10^n}$$

显然 $a_n \leqslant c$. 现在 $a_n + \frac{1}{10^n}$ 与 c 的小数展开式在小数点后,第 n 位小数以前的数字都相同,但 c 的第 n 位小数是 k_n,而 $a_n + \frac{1}{10^n}$ 的第 n 位小数是 $k_n + 1$. 从而

$$a_n \leqslant c \leqslant a_n + \frac{1}{10^n}$$



这些不等式断言 $c \in I_n$,并且因为对每个整数 n 它们都成立,从而 c 在序列的每个区间中,于是在所有区间的交集中.这就证明了序列是正则收缩序列时的定理.

现在令 $I_0, I_1, \cdots$ 是任一收缩序列;令 $I_0 = [a_0, b_0], I_1 = [a_1, b_1], \cdots$,一般,$I_n = [a_n, b_n]$.于是有不等式

$$a_0 \leqslant a_1 \leqslant \cdots \leqslant a_n \leqslant \cdots \leqslant b_n \leqslant \cdots \leqslant b_1 \leqslant b_0$$

选取一个整数 $r \leqslant a_0$ 与一个整数 $s > b_0$,于是得一区间 $I' = [r, s]$,它包含 I_0 与所有其他的区间.用 r 与 s 之间的整数,把区间 I' 分成长为 1 的子区间.令 $I'_0 = [m, m+1]$是这样一个子区间,它的起点 m 前面只有有限多个a_n(收缩序列的区间的起点),而所有的 a_n 都在它的终点$m+1$ 前面.换句话说,那些确实包含序列的 a_n 的I_0的子区间中,I'_0是最靠右的那一个.现在把 I'_0分成十等份,而令 $I'_1 = [c_1, d_1]$ 为 I'_0 的这样一个子区间,使得仅有有限多个 a_n 在c_1 前面,而所有的 a_n 在 a_1 前面.再把 I'_1 分成十等份,而令 $I'_2 = [c_2, d_2]$ 为这样一份,使得仅有有限多个 a_n 在c_2 前面,而所有的 a_n 在 d_2 前面.继续用这个方式以至于对每个 n, $I'_n = [c_n, d_n]$ 是 I'_{n-1} 的十分之一份,并且仅有有限多个 a_n 在c_n 前面,而所有的 a_n 在d_n 前面.因序列 $I'_0, I'_1, \cdots$ 正则收缩,所有这些区间有一个公共点 c.于是,对于所有 n,有 $c_n \leqslant c \leqslant d_n$.因 I'_n 的长度是 10^{-n},有 $c_n \leqslant c \leqslant c_n + 10^{-n} = d_n$,所以

$$d_n \leqslant c + 10^{-n} \quad 与 \quad c - 10^{-n} \leqslant c_n$$

我们要证明:$a_n \leqslant c \leqslant b_n$ 对于每个n 都成立.用反证法.设对于某个整数 N 有 $c < a_N$,因 10 的指数无限制地增加,我们总能找到整数 n,使

$$10^n > \frac{1}{a_N - c}$$

在这些 n 之中,选取一个大于 N 的整数n;对这样一个 n,有

$$a_n \geqslant a_N \quad 与 \quad 10^n > \frac{1}{a_N - c}$$

第二个不等式能改写成 $a_N - c > 10^{-n}$ 或$a_N > c + 10^{-n}$;它与第一个不等式合并,就给出 $a_n > c + 10^{-n}$.因 d_n 不超过$c + 10^{-n}$,从而 $d_n < a_n$,与所有 a_n 在d_n 前面这事实相矛盾.这个矛盾表明对于所有的 n, $a_n \leqslant c$.要证明对于所有的 n, $c \leqslant b_n$,还用反证法.设对于某个整数 N,有 $b_N < c$.然后 $c - b_N > 0$,并且能选取一个大于 N 的整数n 使

$$10^n > \frac{1}{c - b_N}$$

然后有 $b_n \leqslant b_N$ 与 $10^{-n} < c - b_N$.合并这些就给出 $b_n < c - 10^{-n}$.这个不等式与上面关于 c_n 的那一个不等式(即 $c - 10^{-n} \leqslant c_n$)蕴涵着 $b_n < c_n$.因所有 a_n 在

所有 b_n 的前面，这个 $b_n < c_n$ 说明所有 a_n 在 c_n 前面．这是一个矛盾，从而得 $c \leqslant b_n$ 对于所有的 n．这证明了 $c \in I_n$ 对于所有的 n，完成了定理的证明．

6 紧致性

如果 $\mathbf{R}^m$ 的一个子集 X 被包含在某个足够大的球中；也就是，如果有一个点 x_0 与一个数 $r > 0$ 使 $X \subset N(x_0, r)$，我们就说 X 是有界的．线段、圆、球、三角形等等都是有界集的例子．线、半线、射线、平面、$\mathbf{R}^2$ 中圆的外部、全部空间 $\mathbf{R}^m$ 以及有理数都是无界集的例子．直观地说，一个集合，如果沿着它能够一直走向无穷，这个集合就是无界的．

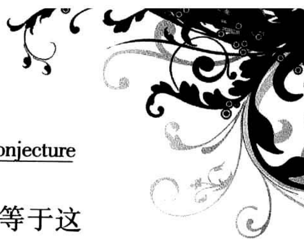
$\mathbf{R}^m$ 的任一子集 X，如果它既是 $\mathbf{R}^m$ 的闭子集又是有界子集，它就具有下述最重要而又值得注意的性质：对于任何连续映射 $f: X \to \mathbf{R}^m$，象集 fX 也闭而有界．本节主要目的就是证明这个事实．这证明有点曲折，不是那么直接；它的历史发展开始于 Cauchy(1789—1857) 的工作，经历过许多阶段．特别是，它蕴涵着 Bolzano-Weierstrass 定理与 Heine-Borel 定理这些分析学中常引用的命题．

在将要给出的证明中，我们首先要表明闭而有界这性质等价于另一个叫做“紧致”的性质．这是论证的主要部分．一旦这个完成了，就容易表明：只要 X 紧致而且 f 连续，fX 就紧致．我们将通过展示无界集与非闭集的一个共同性质，来引进紧致的定义．

令 X 为 $\mathbf{R}^m$ 的一个无界子集，而 x_0 为 $\mathbf{R}^m$ 的一点．想象下面这邻域 $N(x_0, r)$ 的序列，其中 r 取 $r = 1,2,3,\cdots$．这些邻域形成开集的一个扩张序列，其并集是整个 $\mathbf{R}^m$，因为对于每一个 $x \in \mathbf{R}^m$，其距离 $d(x, x_0)$ 小于某个足够大的整数 r．因此，交集 $X \cap N(x_0, r)$，$r = 1,2,3,\cdots$，形成 X 的开集的一个扩张序列，它们的并集是整个 X；但 X 不等于这些开集的任一个，因为 X 无界．此外，X 不包含在任何有限个这些开集的并集之中，因为它们的并集正好是最大的那一个．

现在令 X 为 $\mathbf{R}^m$ 的一个有界但非闭的子集（这种 X 的一个特例是 $[0,1)$．—— 译者），则在 X 的补集 $\mathbf{R}^m - X$ 中有某个点 y 使得每一个邻域 $N(y, r)$ 包含 X 的点（看第 4 节中闭集的定义）．（并看特例 $X = [0,1)$，从而 $y = 1$，—— 译者）对于每个整数 $k = 1,2,3,\cdots$，令 U_k 为以 y 为中心以 $\frac{1}{k}$ 为半径的圆的外部．每个 U_k 是 $\mathbf{R}^m$ 的一个开集，因为如果 $x \in U_k$，则 $N\left(x, d(x,y) - \frac{1}{k}\right)$ 是 x 的一个邻域被包含在 U_k 中．U_k 形成一个扩张的序列 $U_1 \subset U_2 \subset \cdots$，而且它们的并集是 y 的补集，因为对于每个点 $x \neq y$，存在着一个 k 使得 $\left(\frac{1}{k}\right) < d(x, y)$．从而交集 $X \cap U_k$（在特例：$X = [0,1)$ 与 $y = \{1\}$ 时，$X \cap U_k = \left[0, \frac{1}{k}\right)$．——





译者）形成 X 的开集的一个扩张序列，它们的并集是整个 X；但 X 并不等于这些集合的任一个，因为每一个 $N(y,\frac{1}{k})$ 包含 X 的点．此外，X 不在有限个这些集合的并集中，因为它们的并集正好是最大的那一个．

于是，如果 X 是无界的或非闭的，我们能在 X 中找到 X 的开集的一个扩张序列，其并集为 X；但 X 却不是其中任意有限个开集的并集．这就引导出紧致性的定义．不过首先，我们需要"开覆盖"的定义．

定义　令 X 为 $\mathbf{R}^m$ 的一个子集．如果 $\mathbf{R}^m$ 的一些子集的并集包含 X，这些子集所形成的族 C，叫做 X 的一个覆盖；就是说，X 的每一点至少在 C 的一个元素中．X 的一个覆盖 C 叫做有限的，如果 C 的元素的个数有限．X 的一个覆盖 C 叫做包含 X 的一个覆盖 D：如果 D 的每一元素也是 C 的一个元素．（D 也叫做 C 的一个子覆盖．——译者）X 的一个覆盖叫做 X 的一个开覆盖：如果覆盖的每一元素是 X 的（注意：这里的"X 的"这两个字！——译者）一个开集．最后，空间 X 叫做紧致的，如果 X 的每一个开覆盖包含 X 的一个有限覆盖；那就是说，能从 X 的开集的任何无穷族（其并集是 X），挑选出一个有限的子族，其并集也是 X.

如果 X 在 $\mathbf{R}^m$ 中是无界的或非闭的，上面作出的开集的扩张序列是 X 的一个开覆盖．它们的任意有限个的并集就是其中最大的那一个．因为它们中没有一个是 X 的全部，从而它们之中任何有限个都不能覆盖住 X；因而 X 不是紧致的．现在把已得的这些结果（即 $\mathbf{R}^m$ 的一个无界的或非闭的子集必然不是紧致的．——译者），正面地叙述成下面的定理：

定理 6.1　$\mathbf{R}^m$ 的每一个紧致的子集在 $\mathbf{R}^m$ 中是有界的并且是闭的．

我们最终还必须证明逆定理：$\mathbf{R}^m$ 的每个闭的并且有界的子集是紧致的．这是比较困难的，要分几步来完成．（目前，只有那些具有有限个点的集显然是紧致的，因为能从它的任一覆盖选取包含每个点的一个元素）第一步，区间是第一个但并非不重要的例子：

实数的任何一个闭区间 $I=[a,b]$ 是紧致的．

用反证法，假设 C 是 I 的一个开覆盖，而它没有有限的子覆盖成为 I 的一个覆盖，然后引出矛盾．在这个假设下我们要作区间的一个收缩序列 $I=I_0\supset I_1\supset I_2\supset\cdots$ 使每个区间是它前一区间的一半，并且这些区间中没有一个被 C 的一个有限子覆盖所覆盖．然后应用第 5 节中的完全性定理，表明这些区间中确实有一个 I_k，被包含在 C 的一个单独的元素覆盖着，这是个矛盾．

为作收缩序列，我们注意到，根据假设，$I_0=I$ 不被 C 有限地覆盖着．I_0 的中点把 I_0 分成两个闭区间 I'_0 与 I''_0，它们的并集是 I_0．I'_0，I''_0 中至少有一个不被 C 有限地覆盖着；因为，如果不然，C 包含 I'_0 的有限子覆盖 C' 与 I''_0 的 C''，那么 $C'\cup C''$ 会是 I_0 的有限覆盖．选取那不被 C 有限地覆盖着的 I_0 的一半，并且

叫它 I_1.(在两个一半都不被有限地覆盖着时,选取右边的一半,使选择确定)现在把 I_1 分成两半,并照前进行.设 $I_0, I_1, \cdots, I_{k-1}$ 已正式作好,我们像上面那样论证,因为 I_{k-1} 不曾被 C 有限地覆盖着,所以至少它的一半不被 C 有限地覆盖着.取这一半并叫它 I_k.这就完成了存在收缩序列的归纳证明.

由于 $\mathbf{R}$ 的完全性(见第5节),存在着一个点 x 使 $x \in I_k$ 对于所有的 k.因 $x \in I$ 并且 C 覆盖着 I,存在着开覆盖 C 的一个开集 U,使 $x \in U$.因此,存在着一个数 $r > 0$ 使 $N(x, r, I) \subset U$.现在区间

$$I_0, I_1, \cdots, I_k, \cdots$$

包含 x 并有递减的长度

$$(b-a), (b-a)/2, \cdots, (b-a)/2^k, \cdots$$

如果选取一个足够大的 k 使 $(b-a)/2^k < r$,那么 I_k 会完全位于 $N(x, r, I)$ 中.这里出现了矛盾:存在着一个 k 使得

$$I_k \subset N(x, r, I) \subset U$$

所以 I_k 是被 C 的一个单独元素 U 覆盖着,可是我们序列的每一个区间又不以 C 的任何有限子覆盖为覆盖,这个矛盾表明了 I 是紧致的.

在开始下一个例子以前,需要定义 $\mathbf{R}^m$ 的 m 维盒子 B.设对于 $i = 1, 2, \cdots, m$,有数对 $a_i < b_i$,并且限制 $x_i \in [a_i, b_i]$,即 $a_i \leqslant x_i \leqslant b_i$;$B$ 就是 $\mathbf{R}^m$ 中所有以这样的 $(x_1, x_2, \cdots, x_m)$ 为坐标的点 x 组成的一个子集.在 $m = 1$ 时,B 正好是一个闭区间.当 $m = 2$ 时,B 是一个边与坐标轴平行的长方形以及它的内部.当 $m = 3$ 时,B 是一个面与坐标平面平行的长方盒子以及它的内部.

我们也需要能将这盒子 B 再分成较小的盒子.这可以这样做:把区间 $[a_i, b_i]$,$i = 1, 2, \cdots, m$ 的每一个用它的中点 c_i 分成两个区间;所分成的每个盒子以 $[a_i, c_i]$ 或 $[c_i, b_i]$ 作为它的第 i 个区间.于是,当 $m = 2$ 时长方形被直线 $x_1 = c_1$ 与 $x_2 = c_2$ 分成 $4 = 2^2$(个)全等长方形,它的各边长是 B 的对应边长的一半(图6.1).当 $m = 3$ 时盒子被三个平面 $x_1 = c_1, x_2 = c_2$ 与 $x_3 = c_3$ 分成 $8 = 2^3$(个)全等的盒子,它的各条棱长是 B 的对应棱长的一半.一般,B 被 m 个超平面 $x_i = c_i (i = 1, 2, \cdots, m)$ 分成 2^m 个全等的盒子,它的各条棱长是 B 的相应棱长的一半.

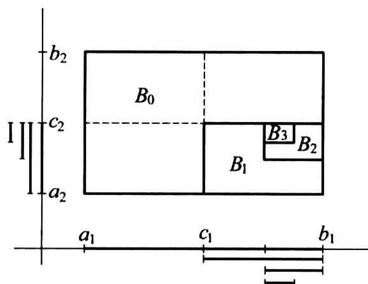


图 6.1

任何 m 维盒子 B 是紧致的.

这里的证明将类似于证明一维盒子(区间)时的方式,用反证法.假设 C 是 B 的一个开覆盖,而它没有有限的子覆盖成为 B 的覆盖.我们来作盒子的一个



收缩序列 $B_0, B_1, \cdots, B_k, \cdots$ 使得 $B_0 = B$，其中没有一个被 C 有限地覆盖着，并且对于每一个 $k > 0$，B_k 是 B_{k-1} 所分成 2^m 个盒子中的一个．由假设，$B_0 = B$ 不被 C 有限地覆盖着．从此开始我们序列的归纳作法．假设 $B_0, B_1, \cdots, B_{k-1}$ 已妥当地选定，考虑 B_{k-1} 所分成的 2^m 个盒子．如果每一个会被 C 的一个有限的子覆盖覆盖着，那么这些 2^m 个子覆盖将会合并为 C 的一个有限的子覆盖，它成为 B_{k-1} 的一个覆盖．因为这不可能，B_{k-1} 的这些盒子中至少有一个不被 C 有限地覆盖着．取 B_k 是这样一个盒子．这就完成了存在着序列 $B_0, B_1, \cdots, B_k, \cdots$ 的归纳证明．(图 6.1 说明 $m = 2$ 时的前三步)

我们断言：对于每个 $k = 1, 2, \cdots$，存在着一点 $x \in B_k$．为了说明这个，考虑盒序列在第 i 坐标轴上的投影，其中 $i = 1, 2, \cdots, m$．在每个轴上的投影组成区间的一个收缩序列．令 x_i 为在第 i 坐标轴上由投影形成的所有区间的一个共同的数．于是，对于所有的 k，$\mathbf{R}^m$ 中坐标是 $(x_1, x_2, \cdots, x_m)$ 的点 x 是 B_k 的一个点．因为 $x \in B$，存在着覆盖 C 的一个开集 U 使得 $x \in U$．因而有一个数 $r > 0$ 使得 $N(x, r, B) \subset U$．令 d 表示 B 的最长棱边的长度．因为在作序列的每一步骤中所有的棱都被平分，从而 B_k 的最长棱的长度是 $\frac{d}{2^k}$．由勾股定理，B_k 的对角线长度最多是 $\frac{\sqrt{m}d}{2^k}$．取一个足够大的整数 k，使

$$2^k > \frac{\sqrt{m}d}{r}$$

于是

$$\frac{\sqrt{m}d}{2^k} < r$$

从而有

$$B_k \subset N(x, r, B) \subset U$$

因此 B_k 被包含在 C 的一个元素中，而这与 B_k 不为 C 有限地覆盖着这一事实相矛盾．我们的假设 B 是不紧致的已经引到一个矛盾；所以 B 是紧致的．

利用下面这一有用的命题，可将能被证明为紧致的集的族大为扩大．

定理 6.2　如果 X 是一个紧致空间(即 $\mathbf{R}^n$ 的紧子集．—— 译者)B 的一个闭子集，则 X 是紧致的．

为证明这定理，取 X 的任一开覆盖 C，并扩大 C 的每个元素使得扩大的元素组成 B 的一个开覆盖 C'．然后，利用 B 的紧致性从 C' 挑选一有限的子覆盖，并且看出那相应的未扩大的元素组成所寻找的 X 的有限的覆盖．

对于 X 的开覆盖 C 的每一个元素 U，令 $U' = U \cup (B - X)$，又令 C' 表示以这些较大的 U' 为元素的族．首先我们要表明 U' 是 B 的一个开集．一点 $x \in U'$，或在 U 中或在 $B - X$ 中．如果 $x \in B - X$，X 在 B 中是闭集的假设告诉我们：存在着一个 $r > 0$ 使得 $N(x, r, B) \subset (B - X) \subset U'$．如果 $x \in U$，在 X 中 U 是开集这事实意味着存在一个 $r > 0$ 使得 $N(x, r, X) \subset U$，因此 $N(x, r, B) \subset$

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

$U \bigcup (B - X) = U'$.这就证明了 U' 在 B 中是开集.

现在令 y 为 B 的任一点;或 $y \in X$,或 $y \in B - X$,如果 $y \in X$,则对于某个 $U \in C, y \in U$,从而对于相应的 $U' \in C', y \in U'$.如果 $y \in B - X$,则对于所有的 $U' \in C'$,有 $y \in U'$.所以 C' 是 B 的一个开覆盖.因为 B 是紧致的,C' 的有限个元素,譬如 $U'_1, U'_2, \cdots, U'_k$ 覆盖 B,因此也覆盖 X.从而 C 的相应元素,即 $U_1, U_2, \cdots, U_k$,组成 X 的一个有限覆盖;对于被 U'_j 覆盖着的 X 的每一个点也被 U_j 覆盖着.这就完成了定理的证明.

把闭的并且有界的集合放在一个 m 维盒子(它已被证明是紧致的)中,然后运用定理6.2,就能证明定理6.1的逆定理.

定理6.3 $\mathbf{R}^m$ 的每个闭的并且有界的子集是紧致的.

令 X 在 $\mathbf{R}^m$ 中为闭且有界.因 X 有界,就有一个点 $b \in \mathbf{R}^m$ 与一个数 $r > 0$ 使得 $X \subset N(b, r)$.令 B 表示中心在 b 且所有棱的长都等于 $2r$ 的 m 维盒子;确切地说,如果 $\mathbf{R}^m$ 中一点 y 的坐标 $(y_1, \cdots, y_m)$ 满足

$$b_i - r \leqslant y_i \leqslant b_i + r, i = 1, \cdots, m$$

y 就在 B 中.于是 B 包含 $N(b, r)$,所以 $B \supset X$;因此 $X \cap B = X$.而且 X 是 B 的闭集;因 X 是 $\mathbf{R}^m$ 的闭集,根据定理4.4′,B 的闭集族就是 B 自身与 $\mathbf{R}^m$ 的闭集族的交集族.X 是紧致的这所求的结论,现在是前面定理的一个推论.

对于 $\mathbf{R}^m$ 的子集,紧致的性质等价于既闭而且有界的性质,现在已完全证明了.这就已准备好来证明本节的主要命题.

定理6.4 如果 X 是一个紧致空间,并且 $f: X \to Y$ 是连续的,则象 fX 是紧致的.

证明 令 C 为 fX 的一个开覆盖.我们需要表明 C 包含 fX 的一有限覆盖.对于每个 $U \in C$,考虑原象 $f^{-1}U$,又令 C' 为全体这些原象所形成的族.因为 f 连续并且 U 是 fX 的一个开集,每个原象 $f^{-1}U$ 是 X 的一个开集.对于每个 $x \in X$,fx 在某个 $U \in C$ 中,因为 C 覆盖 fX;所以 x 在对应的 $f^{-1}U$ 中.于是 C' 是 X 的一个开覆盖.因 X 是紧致的,存在着 C' 的一有限子族 D' 覆盖着 X.C 的对应子族 D 是有限的并且覆盖着 fX;因为,如果 $x \in f^{-1}U$ 并且 $f^{-1}U$ 在 D' 中,那么 $fx \in U$,这里 U 在 D 中.于是 C 包含 fX 的一有限覆盖.这就完成了 fX 是紧致的证明.

下面是定理6.4的直接推论.

推论 如果 X 是 $\mathbf{R}^m$ 的一个闭的并且有界的集合,又如果 $f: X \to \mathbf{R}^m$ 连续,则 fX 是 $\mathbf{R}^m$ 的一个闭的并且有界的集合.

定理6.5 实数的一个紧致的非空集合 X 有一极大值与一极小值;就是说,X 中存在着数 m 与 M,使得 m 是 X 中的最小数,M 是 X 中的最大数.

为着深刻理解这结论的威力,先看一些例子.首先,全体实数集合 $\mathbf{R}$ 没有最





大数与最小数.实际上,数的任意无界集 Y 一定不会有一个极大值或一个极小值;因为如果两个它都有,那么任意兼有这极大值与这极小值的开区间会包含 Y 的全部,于是 Y 将会是有界的.其次,还存在着既无一个极大值也无一个极小值的有界集.例如一个开区间 (a,b) 就是既没有一个最大数也没有一个最小数.这些例子表明:要获得定理的结论,我们必须既要求 X 是有界的,并且还必须增加某些不为开区间所满足的额外条件.因为一个紧致的集合是既闭并且有界,单单这紧致性的条件就保证了有界,并且排斥了开区间.

让我们继续证明定理.因为 X 是紧致的,它就有界;因而存在着一个包含 X 的闭区间 $I_0=[a_0,b_0]$.我们来造一个具有下列性质的区间的收缩序列 I_0, $I_1,\cdots,I_k,\cdots$:每个区间 I_k 是 I_{k-1} 的一半;每个 I_k 至少包含 X 的一个点;最后,I_k 的右端点 b_k 是 X 的一个上界,就是说,对于所有的 $x\in X$,我们有 $x\leqslant b_k$,$k=0,1,2,\cdots$.毫无疑问 I_0 含有 X 的一个点(因 X 是非空的),并且 b_0 是 X 的一个上界.假定已妥当地选取了 $I_0,I_1,\cdots,I_{k-1}$.令 c 为 $I_{k-1}=[a_{k-1},b_{k-1}]$ 的中点.如果 c 是 X 的一个上界,取 $I_k=[a_{k-1},c]$,如果 c 不是一个上界,取 $I_k=[c,b_{k-1}]$.在两种情形的任一种时,I_k 都具有所要求的性质.

由于 $\mathbf{R}$ 的完全性(见第 5 节),存在着一个数 M,使得对于每个 $k=0,1,2,\cdots$,有 $M\in I_k$.我们将首先证明 M 属于 X.设不然,因 X 是闭集,补集 $\mathbf{R}-X$ 就是开集;于是会有一个 $r>0$ 使得 $N(M,r)\subset\mathbf{R}-X$.如果 d 表示 I_0 的长度,则 I_k 的长度是 $\frac{d}{2^k}$.对于一个充分大的整数 k,我们有 $\frac{d}{2^k}<r$,所以 I_k 包含 M 并且它的长度小于 r;因而

$$I_k\subset N(M,r)\subset\mathbf{R}-X$$

这与每个 I_k 含有 X 的一个点的事实相矛盾.从而 M 必须属于 X.

现在我们证明:M 是 X 的最大数.假设不然,就会有一个 $x\in X$ 使得 $x>M$.取 $r=x-M$,因而 $r>0$,选取一个整数 k,使其数值足能使 $\frac{d}{2^k}<r$.因为 $M\in I_k$ 并且 I_k 的长度小于 r,从而 $b_k<x$;所以 b_k 不是 X 的上界.但在造区间时 b_k 是一个上界.这个矛盾证明了 M 是 X 的最大数.

存在着极小值的证明同样地进行.区间的收缩序列是这样选取的,使每个区间包含 X 的某个点,并且每个区间的左端点是 X 的下界.证明的细节留给读者.运用由 $fx=-x$ 所定义的映射 $f:\mathbf{R}\to\mathbf{R}$,从极大值的存在也能得出极小值的存在.因为 X 是紧致的,定理 6.4 断言 fX 是紧致的.然后 fX 有一极大值,譬如 M'.从而 fM' 是所求的 X 的极小值.

本节最后的定理提供了第 1 节中主要定理的部分结论.

定理 6.6 如果 X 是 $\mathbf{R}^m$ 的一个闭的、有界的非空子集,又如果 $f:X\to\mathbf{R}$ 是

定义在 X 上的一个连续的实值函数,则象 fX 有一极大值 M 与一极小值 m.

因 X 既闭并且有界,它就是紧致的.因 X 是紧致的并且 f 连续,象 fX 就是紧致的.因 fX 是紧致的又是实数的非空集合,前面的定理保证了 m 与 M 的存在.这就完成了证明.

7 连通性

为了证明本文的主要定理,我们需要闭区间的两个重要拓扑性质.这两个的第一个紧致性,已在第6节中研究了.现在讨论另一个叫做"连通性"的性质.

有些空间能自然地分成两部分或更多的部分.例如,两不相交直线所组成的空间能分成两条直线.作为另一个例子,平面中圆的补集有两部分:圆的内部与圆的外部.再者,如果 p 是直线 L 的一点,那么 L 中 p 的补集自然是由 p 决定的两个半直线(p 的挖去把 L 切割成两部分).

在上面所举的每一个例中,恰出现一种自然的分法.有理数集合 **Q** 能有许多种分法.每一个无理数 x 产生 **Q** 的一种分法:大于 x 的那些有理数的部分与小于 x 的那些有理数的部分.无理数集合能被每一个有理数用同样方法分开.

另一方面,某些集合不能以任何自然方式分成一些部分,这是真的,例如,一直线、一线段、一平面与一圆片.

当然强行分成几份是可能的.例如,如果 I 是区间 $[a,b]$ 并且如果 c 是一个数使得 $a < c < b$,则 c 把 I 分成两个区间 $[a,c]$ 与 $[c,b]$.不过,因它们都有共同的 c,我们不把这种分法看成一种妥当的分法.从这两个分开的集合之一删去 c,譬如说从第二个删去 c,我们就获得一个妥当的分法.令 $A=[a,c]$ 与 $B=(a,b]$,则 $A\cup B=I$ 与 $A\cap B=\varnothing$.我们不认为 I 的这种分法或"分裂"是自然的.因为集合 B 在点 c 处"接触"到集合 A.如果我们也从 A 删去 c 来避免这个"接触性",那么 $A\cup B\neq I$,而 $A\cup B$ 是 I 中 c 的补集;因而这个例子类似于直线 L 中一点 p 的补集的例子.

我们需要的正确概念,现在叙述于下:

定义 空间 X 的一个分离是 X 的一对非空子集 A 与 B,使 $A\cup B=X$,$A\cap B=\varnothing$,并且 A 与 B 都是 X 的开集.没有分离的空间叫做连通的.

例如,考虑平面中一个圆 C 的补集 X.令 A 为圆的内部,B 为圆的外部;就是说,A 包含与圆心距离小于圆半径的所有点,而 B 是 X 中 A 的补集.一个分离的条件易于验证.X 中 A 与 B 都是开集这一事实,从图7.1中是显而易见的;A 的每一点有包含在 A 中的一邻域,而 B 的每一点有包含在 B 中的一邻域.

令 L 为一直线,p 为 L 的一点,而 X 为 L 中 p 的补集.令 A 为 L 上位于 p 左边的点的集合(图7.2),B 为 L 上位于 p 右边的点的集合.此外 A 的每个点有一



邻域在 A 中,同样,B 的每个点有一邻域在 B 中.

现在回忆区间 $I = [a,b]$ 的"强行"分为 $A = [a,c]$ 与 $B = (c,b]$ 的分隔.如果用分离的三个条件来检验,我们发现除下面一个外都满足.A 在 I 中不是开集,因为点 $c \in A$ 没有邻域完全在 A 中.

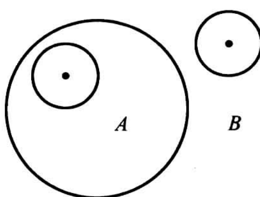


图 7.1

这些预备性的考虑说明,分离的定义与连通空间的定义精确地表达出我们心目中粗糙的几何概念,后面的定理将完全验证这些定义的合理性.

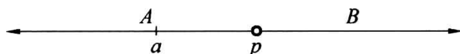


图 7.2

分离的定义能改用若干等价的提法.因为 A 与 B 在 X 中互为补集,一个是开集当且仅当另一个是闭集.于是能同样地要求 A 与 B 都是 X 中的闭集.我们也可以不管 B,而说 X 的一个分离是 X 的一个子集 A,它既是 X 的一个开集又是 X 的一个闭集,并且它既不是空集又不是 X.从而它在 X 中的补集 B 有同样性质.(记住,$\varnothing$ 与 X 在 X 中都既是开的又是闭的)

这样,下面的任何一个都可以作为空间 X 的一个分离 A 与 B 的定义:

1. A 与 B 都是 X 的非空子集使得 $A \cup B = X, A \cap B = \varnothing$,并且 A 与 B 在 X 中都是开的;

2. A 与 B 都是 X 的非空子集使得 $A \cup B = X, A \cap B = \varnothing$,并且 A 与 B 在 X 中都是闭的;

3. A 是 X 的一个子集,它既是 X 的一个开集又是 X 的一个闭集,并且它既不是 $\varnothing$ 也不是 X.

证明一个空间不连通,往往比证明它连通容易.在第一种情形,我们只需要举出一个分离,并且验证它确是一个分离,而在第二种情形,则必须顾到 X 的除 $\varnothing$ 与 X 以外的所有开集,即证明每一个这样的开集不是 X 中的闭集.下列定理不仅证明某些简单的空间是连通的,并且也提供了验证许多空间连通性的技巧.

定理 7.1 实数的一个闭区间是一个连通集.

令 I 为 $\mathbf{R}$ 的一个闭区间,又令 A 为 I 的一个闭区间,A 既不是 $\varnothing$ 又不是 I.为了证明这定理,我们将证明 A 在 I 中不是开的.因 $A \neq \varnothing$ 又 $A \neq I$,就有一个点 $a \in A$ 与一个点 $b \in I - A$.令 I' 表示区间 $[a,b]$(或 $[b,a]$,如果 $b < a$).为了画出这个,想象区间 I 由图 7.3 所示的子集 A 与 B 组成,其中 A 是闭集而 B 是 A 在 I 中的补集.因 A 与 I' 都是闭集,所以它们的交集 $A \cap I'$ 是闭集.因 $A \cap I'$ 又有界,它是一个紧致的集合.由定理 6.5,$A \cap I'$ 有一极小值 m 与一极大值 M.

如果 $a < b$,则 $m = a$,因 a 是I' 的左端点.因右端点 b 不在A 中,我们有 $m = a \leqslant M < b$.从而 M 的每个邻域包含介于M 与b 之间的数;这些都不在 A 中,因此 A 不是开的.如果 $a > b$,则 $M = a, b < m \leqslant a$,并且 m 的每个邻域包含介于 b 与m 之间的数;这些都不在 A 中,所以 A 仍不是开的.这就证明了 I 的唯一的既开又闭的子集是 I 与$\varnothing$.所以 I 是连通的.

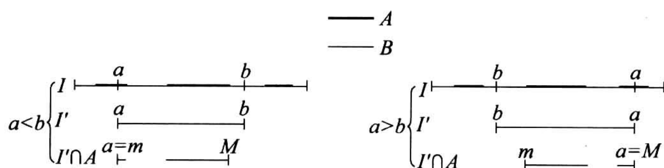


图 7.3

定理 7.2 如果 $f: X \to Y$ 是一个连续映射,并且 A 与 B 是象fX 的一个分离,则原象 $A' = f^{-1}A$ 与$B' = f^{-1}B$ 形成X 的一个分离.

我们必须验证 A' 与B' 这对集合满足一个分离的各个条件.因为 A 不是空的,就有一个点 $y \in A$.因为 $A \subset fX$,就有一个 $x \in X$ 使$fx = y$;所以 $x \in A'$,因而 A' 不是空的.同样,B' 不是空的.任意点 $x \in X$ 有它的象在$fX = A \cup B$ 中,所以 $fx \in A$ 或$fx \in B$.相应地,$x \in A'$ 或$x \in B'$.这表明 $X = A' \cup B'$.如果 A' 与B' 有一公共点x,那就蕴涵 fx 是A 与B 的公共点.这是不可能的,所以 $A' \cap B' = \varnothing$.最后,因为 A 与B 在fX 中是开的并且f是连续的,它们的原象在 X 中是开的(见定理 4.5).这证明了 A', B' 是 X 的一个分离.

这定理能再说得简短些:如果f连续并且fX 不连通,则 X 不连通.因此一个等价的断言是:

推论 如果 $f: X \to Y$ 是连续的并且 X 是连通的,则 fX 是连通的.

这一提法是对我们最有用的一种;例如,它使我们能证明:

推论 每一条线段是一个连通集.

证明 如果 L 是一条线段,I 是一个区间,就有一个相似变换 $f: I \to L$,使得 $fI = L$.因为相似变换是连续的(见第 3 节)并且 I 是连通的,因此 L 连通.

两条相交于一点的线段合起来,形成一个连通集;这在直观上是显然的.它的证明依赖于下面两个引理.

定理 7.3 如果 X 是不连通的并且A 与B 是X 的一个分离,则 X 的每一个连通子集或者完全在 A 中,或者完全在 B 中.

设 C 是X 的一个子集,它包含 A 的一点与B 的一点,则 $C \cap A$ 与$C \cap B$ 都不是空集,都是 C 的一个开集;它们的并集是 C 而它们的交集是空的.所以 C 是不连通的.这就证明:一个连通子集不能同时包含 A 与 B 的点.

定理 7.4 如果两个连通集合 Z 与W 有一个公共点,则 $Z \cup W$ 连通.





用反证法.设 $Z \cup W$ 有一个分离A,B.令 c 为Z与W的公共点.如 $c \in A$,则 Z 与W都是包含A 的一个点的连通集.根据引理 7.3(连同 $X = Z \cup W$),Z 与 W完全在A中.故 B 必为空的.如 $c \in B$,我们看出 Z 与W完全在B中,于是 $A = \varnothing$.这两种情况的任一点,都出现矛盾;所以 $Z \cup W$ 是连通的.

引理 7.4 的一个简单应用表明:两条相交于一点的线段合起来形成一个连通集;在这相交的两线段上,逐次接上另一条线段,就作成一条折线,所以每一条折线是一个连通集.

下一个定理是证明某些空间具有连通性的重要工具.

定理 7.5 一空间 X 是连通的$\Leftrightarrow X$ 的每一对点位于X的某个连通子集中.

定理的"$\Rightarrow$"这部分的证明是明显的、平凡的.因为整个 X 当然是它自身的一个连通子集,并且包含 X 的每一对点.从而 X 自身就是结论中的"某一连通子集".

为着用反证法来证明定理的"$\Leftarrow$"这另一半,假设 X 有这样的性质:它的每一对点位于 X 的一连通子集中,并设 X 不连通.令 A 与B 为X的一个分离,又令 $x \in A$,$y \in B$(记住 A 与B 都不是空集).由假设,在 X 中存在着包含x 与y 的一个连通子集 C.但是根据引理 7.3,C 或者完全在A中,或者完全在 B 中.这个矛盾指出 X 不能有分离,所以 X 是连通的,定理得证.

记住 $\mathbf{R}^n$ 的一个子集叫做凸的,如果它包含它的每一对点的连接线段(例如;平面中一个圆的内部是凸的,而它的外部不是).因为线段是连通的,定理 7.5 蕴涵着

推论 每一个凸集是连通的.

一个连通集不一定是凸集.尽管一个圆的外部不是凸集,但它却是连通的,因为它的任何两点都可以用位于外部的一条折线连接起来(图 7.4),同样,介于两圆之间的一个平环的任两点可用一条折线连接;因此,一个平环是连通的.

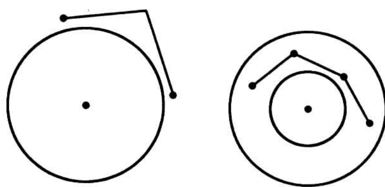


图 7.4

现在我们回转来研究直线的子集(即实数的子集),得到关于连通性的最后命题.

定理 7.6 实数的一个紧致、连通集合是一个闭区间.

本定理的逆定理,即一闭区间既紧致又连通,已经证明过了(见第 6 节与定理 7.1).

证明 令 X 为实数的任一连通集,又令 a 与b 为X中的数,并且 $a < b$.先证明任何适合 $a < c < b$ 的数c 也在X中.这一点 c 决定它在$\mathbf{R}$中的补集 $\mathbf{R} - c$ 的一个分离;令 A 包含所有小于c 的数,又令 B 包含所有大于c 的数.如果,同

我们所要证明的相反，X 不包含 c，它就会是 $A \cup B$ 的一子集，又因 X 是连通的，根据引理 7.3，它将完全在 A 中或完全在 B 中. 但 X 包含 a 与 b，所以这不可能. 于是我们证明了实数的一连通集包含它的任何两数之间的所有数.

如果 X 还是紧致的，定理 6.5 断言：X 有一极小值 m 与一极大值 M. 从而 X 的确是这闭区间 $[m, M]$.

8 拓扑性质与拓扑等价

这些节的基本任务是证明第一节所述的主要定理：如果实值函数 fx 在闭区间 $a \leqslant x \leqslant b$ 中有定义且连续，则函数 fx 有一极小值，一极大值，并取得介于二极值间的任一值. 我们已经完成为证明所需要的全部工作；只需把各部分汇总起来. 下列三个命题都已证明：

1. 实数的一个闭区间是一紧致的连通集.

2. 紧致集的连续象是紧致的，连通集的连续象是连通的.

3. 实数的一个紧致并连通的集合是一闭区间.

把第 6 节中已证明的关于紧致性的断言，与第 7 节中已证明的关于连通性的断言合并起来，就得出前两个命题的每一个. 第三个命题是定理 7.6. 这三个命题结合在一起断言：一个闭区间在 $\mathbf{R}$ 的连续象本身是一个闭区间. 这恰好是主要定理的另一个说法.

我们已经做了比证明主要定理超出很多的工作；我们已证明了好些具有相当一般性的定理，并且分析了论证，使能获得与主要定理相似的定理不会有更多的困难. 例如，把一个闭且有界集是紧致的这一结论(第 6 节)，跟命题 2 与 3 结合起来，我们就能断定：

如果 X 是 $\mathbf{R}^n$ 中的一个闭的、有界的、并且连通的集合，并且如果 $f: X \to \mathbf{R}$ 连续，则象 fX 是一闭区间.

在 $\mathbf{R}^n$ 的闭的有界连通子集的许多不同类型中，有一种是 $\mathbf{R}^3$ 中一个球的表面. 因此，一个定义在球上的连续实值函数必有一极大值和一极小值，并能取遍其间的任一个值. 我们现在能看出：主要定理中关于 f 的定义域是闭区间这一假设是不必要的限制；只要要求定义域是闭的、有界的、连通的就够了.

现在我们开始回答问题：什么是拓扑学?

定义 $\mathbf{R}^m$ 的子集 X 的一种性质叫做拓扑性质，如果它等价于一个只利用 X 的开集的概念与集合论的一些标准概念(元素、子集，补集、并集、交集、有限集、无限集等)来下定义的性质. 简短地说，$\mathbf{R}^m$ 的子集 X 的拓扑性质就是可以表达为 X 的开集族的性质的那一种性质.

紧致性与连通性都是拓扑性质. 建议读者仔细地复习第 6 节与第 7 节中给





出的这些概念的定义，并且注意其中没有提到的 X 的大小、形状、长度、面积以及体积等性质．同样，X 中一闭集是一拓扑概念，因为根据定义，一闭集是 X 中一开集的补集．

一旦知道了一种性质或概念是拓扑的，我们就可以自由地利用它来定义别的拓扑性质与概念；例如闭集、紧致性以及连通性概念，就可以这样利用．

下面是一些特种集合的拓扑性质的几个例子．一直线 L 是一连通集，而 L 的每一点在 L 中的补集不连通．换句话说：挖去直线的任一点，这直线就不连通．一个圆没有这性质；但是挖去它的任一对点，它就不连通了．一平面的一些拓扑性质是：它连通，它不紧致，挖去它的任意有限个点，它并非不连通．

$\mathbf{R}^m$ 的子集 X 的一种性质如果牵涉到 X 或它的子集的诸如大小、形状、角度、长度、面积或体积等特性，它就不可能是拓扑性质．例如有界性指的是 X 的大小．X 在 $\mathbf{R}^m$ 中为闭集的性质牵涉到 $\mathbf{R}^m - X$ 这集合，而不只是牵涉到 X 的一些开集．从表面看来，有界与闭这两个性质没有一个有可能是拓扑性质．但是马上做轻率的结论，在这里是危险的．如果考虑 X 在 $\mathbf{R}^m$ 中既有界又闭这一性质，我们会同样地认为它不是拓扑性质；但在第 6 节中已证明，它等价于紧致性这一拓扑性质．显然，要肯定某一种性质不是拓扑性质，需要一种实用的检验法．这样一种检验法根据于两个点集拓扑等价这个概念．

定义　$\mathbf{R}^m$ 的子集 X 与 $\mathbf{R}^n$ 的子集 Y 叫做拓扑等价的（或同胚的），如果存在一个一对一的函数 $f: X \to Y$ 使得 f 连续，并且 $f^{-1}: Y \to X$ 也连续．此外，这函数 f 叫做拓扑等价（或同胚）．

我们来考虑一些例子．我们已经观察到任两线段是相似的，并且相似变换是连续的．因相似变换的反函数也是相似变换，从而任两线段是拓扑等价的（图 8.1）．

如图 8.2 所示，我们可以利用以 z 为中心的一个径向射影来定义一线段与一圆弧之间的拓扑等价．作由邻域 $N(fx, \varepsilon)$ 与点 z 所确定的楔形域，然后取正数 δ 如此小使 $N(x, \delta)$ 落在这个楔形域内，就可证明 f 是连续的．因为 f^{-1} 缩短距离，它也连续．

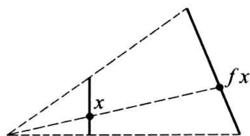


图 8.1

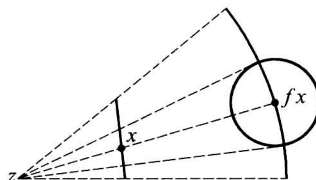


图 8.2

事实上，一条升降频繁的曲线能与一线段拓扑等价．图 8.3 表示定义在区间 I 上的一连续函数 f 的图形 C．令 $g: C \to I$ 为垂直射影，所以 $g(x, fx) = x$．

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

显然 g 是一对一的,而且对于所有 $x \in I, g^{-1}x = (x, fx)$. 作为垂直射影,$g$ 缩短距离,故 g 连续. g^{-1} 的连续性是 f 连续的推论. 所以任何连续函数 f 的图形 C 与线段 I 拓扑等价.

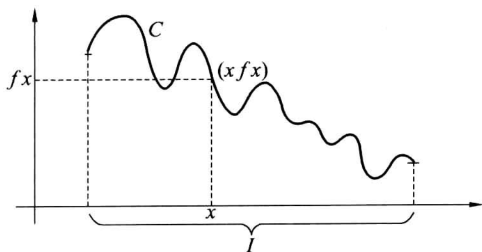


图 8.3

粗略地说,设一质点连续地从一点 p 走到另一点 q,描出一条不自交的曲线,这曲线与一线段拓扑等价(图 8.4).

挖去了北极 p 的一个球面 S,到它赤道平面 P 的球极平面射影提供同胚的另一个例子(图 8.5). 这射影是 $S-p$ 与 P 之间的一个拓扑等价. 如果我们把这射影限制在单独一个通过 p 的大圆 C 上,就得到 $C-p$ 与一直线 L 之间的一个拓扑等价.

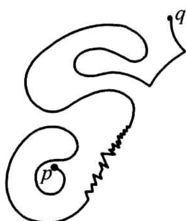


图 8.4

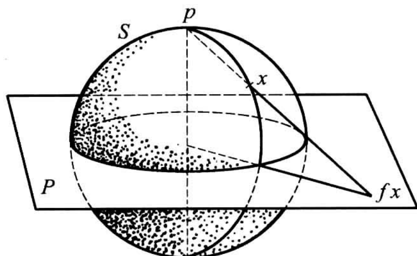


图 8.5

已经说明了拓扑等价概念,让我们返回来讨论拓扑性质. 下面的定理讲明这两个概念的基本关系.

定理 8.1 如果 $\mathbf{R}^m$ 的一个子集 X 与 $\mathbf{R}^n$ 的一个子集 Y 拓扑等价,则 X 与 Y 这两个子集的每一个,有另一个所具有的每一种拓扑性质.

这是显而易见的,因为一个拓扑等价或同胚 $f: X \to Y$ 建立起两个集合的点之间的一一对应关系,与它们的子集之间的一一对应关系(因 $A \subset X$,就有 $fA \subset Y$,又因 $B \subset Y$,就有 $f^{-1}B \subset X$),这对应关系是这样的,它使得开集对应于开集,并保留集合论的关系与运算(例如在 X 中 $A \subset B$ 当且仅当在 Y 中 $fA \subset fB$). 对于有关 X 的点,X 的子集,X 的开集,以及它们的集合论关系的任何真命题,如果把其中 X 的所有点与子集用它们在 Y 中的同胚象替换,所得到的关于 Y 的新命题自然也真.



我们用不连通这一性质来说明这论点.用开集来说,命题是:X 有两个非空开集 A 与 B,使得 $A \cup B = X$ 与 $A \cap B = \varnothing$.如果取 f 作用下的象并利用明显的关系 $fX = Y, f\varnothing = \varnothing, f(A \cup B) = fA \cup fB$ 与 $f(A \cap B) = fA \cap fB$,则得:Y 有两个非空开集 fA 与 fB 使得 $fA \cup fB = Y$ 与 $fA \cap fB = \varnothing$.故 Y 不连通.

我们用定理 8.1 来说明 $\mathbf{R}^m$ 的子集 X 的某些性质不是拓扑性质.因为一线段拓扑等价于任一别的线段,所以线段的长度不是拓扑性质.因为一线段与圆的一段弧等价,所以线段的平直不是拓扑性质.因为在球极平面射影下,挖去一点 p 的球面 S 与一平面 P 等价,所以 $S - p$ 的有界性不是一个拓扑性质.因为 P 在 $\mathbf{R}^3$ 中是闭的,而 $S - p$ 在 $\mathbf{R}^3$ 中不是闭的,P 在 $\mathbf{R}^3$ 中是闭集这一性质不是拓扑性质.

到此,对于"什么是拓扑学"这问题的回答应该是相当明显了:拓扑学是研究点集拓扑性质的学科.

这是一个满意的回答,但是并不完全.我们还必须包括函数的拓扑性质.设 $f: X \to Y$ 是一函数,其中 $X \subset \mathbf{R}^m$ 与 $Y \subset \mathbf{R}^n$.f 的一个性质叫做拓扑性质,如果它等价于一种性质,其定义只用到 X 与 Y 的开集,象与原象,以及集合论的标准概念.

例如,连续性就是函数的一个拓扑性质,因为定理 4.5 说到 f 连续当且仅当 Y 的每个开集的原象是 X 的一个开集.容易找出函数的拓扑性质.另一个例子,f 是一个常函数这一性质是一拓扑性质.再有,对于每一点 $y \in Y, f^{-1}y$ 是 X 的一个紧致子集,$f: X \to Y$ 的这性质是一个拓扑性质.

原问题的完满回答是:拓扑学是研究点集与函数的拓扑性质的学科.

按照这些定义的精神,我们再检查一下主要定理的证明.在本节开始时,主要定理已分成三个命题.第二个命题是纯拓扑的;它讲到:X 的一拓扑性质(紧致性)与 f 的一拓扑性质(连续性)蕴涵 fX 的一拓扑性质(紧致性).用连通性替代这命题中的紧致性,新命题同样成立.第一个命题叙述了 $\mathbf{R}$ 中的闭区间这一熟知对象的两个拓扑性质.第三个命题是第一个的逆命题:紧致性与连通性这两个拓扑性质,刻画出 $\mathbf{R}$ 中闭区间的特征,把闭区间从 $\mathbf{R}$ 的其他子集区别出来.因此我们能得出结论:主要定理的证明几乎全是拓扑的证明.

拓扑学曾经被人们叫做橡皮几何学.如果要描绘与一特定点集 X 拓扑等价的那些点集,把 X 看成是用橡皮制成的,确是一种好的直观办法.如果把橡皮制成的 X,这儿扯长些,那儿缩短些,并且扭弯一些(但决不撕裂开,也不把不同部分粘连起来),变形成点集 Y,则 X 与 Y 是拓扑等价的.例如,一个小球面(气球)能膨胀成一个大球面,它也能被挤压成一椭球,还能再挤压成一个哑铃的表面.还有,能让一个吹鼓了的球面收缩,直到它恰紧绷在一长方盒子或一四面体表面(图 8.6).

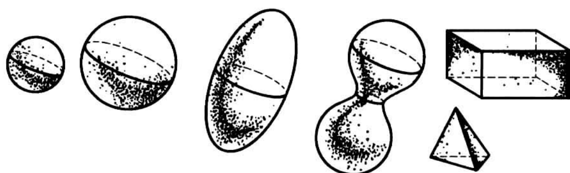


图 8.6

因为两个拓扑等价的点集具有完全相同的拓扑性质,拓扑学者把它们看成本质上是相同的(从拓扑观点来看,无区别的).这类似于欧氏几何学中把两个重合的图形看做完全等价这一观点.所以拓扑学者被人们称为不能区别出一个油炸面圈与一咖啡杯的数学家.图 8.7 表示一个实心油炸面圈变形成一个咖啡杯子的几个中间步骤.



图 8.7

在每一种经典几何学中,都有等价图形这一概念.已经指出,在欧氏几何学中,两图形等价,如果它们重合,特别是,如果有一刚体运动,把一图形变成另一个.在射影几何中,如果有一个射影变换把一图形变成另一个,则这两图形等价.射影变换包括重合与相似,以及足够多的别种变换,因而任何两个三角形是等价的,任一圆与任一椭圆是等价的(图 8.8).

在微分几何中,等价叫做等距(等度量).这里,两个图形等价,如果在它们的点之间存在着一一对应,使得在定义域中任一曲线的长度等于值域中相应曲线的长度.例如,圆柱面的一部分能铺开放在平面的一部分上面.同样,锥面的一部分能铺开放在平环的一扇形上面(图 8.9).所以这些曲面都等距.

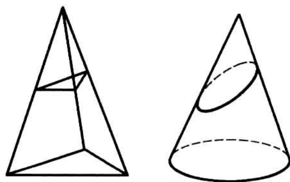


图 8.8

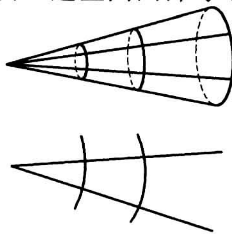


图 8.9

重合性,射影对应性,与等距性都是拓扑等价;因为在每种情形,两个等价图形之间的对应关系是一对一的,并且在每一方向都连续.从而一个这种图形的每一个拓扑性质也是另一个图形的一个性质,所以一个拓扑性质也是欧氏几



何意义下、射影几何意义下与微分几何意义上的一个性质.结果,拓扑学的一个定理自动地成为这些几何学中每一种的一个定理.由此,有充分正当的理由来说:拓扑学是基本的几何学.

9 关于不动点的一个定理

如果一个函数 f 的定义域与值域是同一个集合 X,f 就叫做 X 的一个自映射.X 的自映射 f 可能会使 x 的某个点 X 原地不动,就是可能会使 $fx = x$.具有这性质的点 x 叫做自映射 f 的一个不动点.如果把一圆片绕中心旋转一直角,圆片中心是唯一的不动点.这映射如果限制在圆片的边缘圆周上,就没有不动点.一空间的每一个常值自映射有一个不动点.于是一个集合的一个自映射有否不动点,取决于这集合与这自映射.但是,在集合为一线段(闭区间)的情形,有下面值得注意的结果.

定理 9.1 一线段的每一个自映射至少有一个不动点.

设在直线上引进了坐标,使线段成为一区间 $[a,b]$.于是线段的一个自映射恰好是一个连续函数 $f:[a,b]\to[a,b]$.对于每一个 $x\in[a,b]$,用 $gx = fx - x$ 来定义一个新函数 $g:[a,b]\to\mathbf{R}$.于是 g 给出从 x 到它的象 fx 的有向距离.当 fx 在 x 之右,即 $fx > x$ 时,距离是正的,当 fx 在 x 之左,距离是负的.找 f 的一个不动点,就是找使 g 为零的点.如果两个端点的任一个是不动点,就已符合定理的结论,没什么还要证明的.现在假设两个都不是不动点.因为 fa 与 fb 都在 $[a,b]$ 中,$a < fa$ 而 $fb < b$;因而 $ga > 0$,$gb < 0$.因为 g 连续(它是两连续函数之差),主要定理断言:g 取 ga 与 gb 之间的每一个值.故对于某个 $x\in[a,b]$,必有 $gx = 0$,而这个 x 就是所寻求的 f 的不动点.

设图 9.1 的曲线是 f 的图形.我们可以从图形的观点来检验定理 9.1.f 的不动点就是 f 的图形与对角线的交点(即如果 $fx = x$,则 $(x,fx) = (x,x)$ 在对角线上).因为 $a < fa$,所以 (a,fa) 这点位于对角线上方,同样 (b,fb) 这点位于对角线下方.因为对角线把平面分成它上方与它下方两部分,而 f 的图形是一连通集,所以它必与对角线相交.函数 g 给出图形与对角线之间的垂直距离.

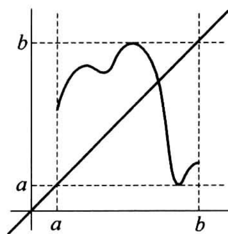


图 9.1

10 圆到直线的映射

圆具有下述的一个奇特性质:

定理 10.1 圆到直线的每一个映射都把某对对径点映成同一个象点.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

证明　令 $f: C \to L$ 为圆到直线的一个映射. 在 L 上引进坐标,就可把 f 的值域看做是实数 $\mathbf{R}$. 考虑 C 上的一对对径点 p 与 p'(图 10.1); 设它们在 L 上的象点分别有坐标 $fp = a$ 与 $fp' = b$. 这一对固定的对径点 p 与 p' 把圆分成两个闭的半圆,都以 p 与 p' 为端点. 令 x 与 x' 为圆上的变动的一对对径点;当 x 沿着一个闭的半圆从 p 走到 p', x' 就沿着另一个闭的半圆从 p' 走到 p. 现在观察由下面的式子定义的函数 g

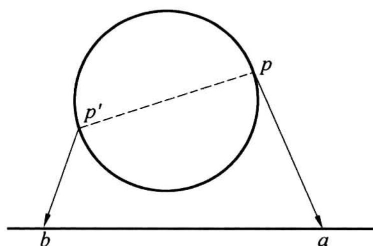


图 10.1

$$gx = fx - fx'$$

它的定义域是 x 走过的闭的半圆;函数 g 连续,因为 f 连续. 显然,在定义域的两端处

$$gp = fp - fp' = a - b, gp' = fp' - fp = -(a - b)$$

因为 $a - b$ 或者等于零,或者不等于零,所以相应地,函数 g 在 p 与 p' 处或者都是零(这时, p 与 p' 的 f 象相同),或者 g 在 p 与 p' 处的值有相反的符号. 在第二种情形,应用主要定理到 g 的半圆定义域,就获得其中的一点 q 使 $gq = 0 = fq - fq'$. 因而 $fq = fq'$,即一对对径点 q 与 q' 的 f 象相同.

类似于圆的对径点的,在椭圆上是对极(Antipodal)点;椭圆的一对对极点关于椭圆中心对称. 因为圆是椭圆的特例,所以对径点是对极点的特例. 因而自然会问:对于椭圆的对极,是否成立类似的定理.

如果圆 X 与椭圆 Y 有同一个中心 z,并且在同一平面上,下述的作图最容易给出一个同胚. 作以 z 为起点的任一射线交 X 于一点,也交 Y 于一点;这两个点的对应就是一个同胚. 这实质上是第 3 节中所提到的从 $\mathbf{R}^2$ 到 X 的满的径向射影,并且在该节中已经证明它是连续的. 如果 X 与 Y 不是同心的,则先把 Y 径向射影到与 Y 同心的圆 X' 上,然后再找同胚(图 10.2). 因为 X 与 X' 相似, X 就与 X' 拓扑等价,而 X' 已与 Y 拓扑等价. 此外,由径向射影与相似变换复合而成的同胚 $Y \to X$,保持对极性;即如果 q 与 q' 是对极点,并且 q 的象是 p,则 q' 的象是 p 的对径点 p'. 所以把对极点与对径点理解为扮演同一角色时,这对径点定理对于椭圆也成立.

类似的论点还证明,对于任意星形闭曲线,如图 10.3 中的多边形 B,这结果也成立. 通过一个以 z 为中心的,从 B 到圆 C 的满的径向射影,就获得一个同胚,它把过 z 的同一直线上的 B 的每一对点变成该直线上 C 的一对对径点.



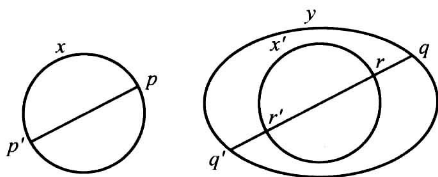


图 10.2

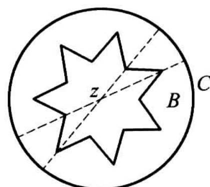


图 10.3

11 薄煎饼问题

薄煎饼的第一个问题可以大致叙述如下:设想两个形状不规则的薄煎饼放在同一个圆形的案板上.试证:一刀能把两个饼正好都切成两半.如果每个薄煎饼的形状都恰好是圆形的,那么沿着它们的圆心的连线切一刀,就能达到目的.但是在薄煎饼的形状不加限制时,问题就变得比较困难.下面是精确的数学定理.

定理 11.1 如果 A 与 B 是同一平面中的两个有界区域,那么平面中存在一条直线,它把每个区域分成等面积的两半.

所谓平面中一个区域,指的是平面的一个连通开子集.定理即使在两个薄煎饼之一堆迭在另一个上面时仍旧成立,也就是说,当两个区域部分地重迭起来时仍旧成立.

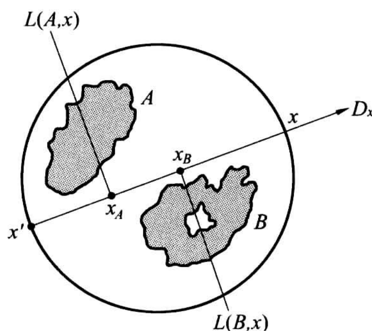


图 11.1

因为证明相当长,我们先给出它的主要步骤,而把两个次要命题的证明推迟.因为 A 与 B 是有界的,我们可以选取一个圆 C 把 $A \cup B$ 包含在它的内部(图 11.1).令 z 与 r 分别表示C的中心与半径.对于任一 $x \in C$,令 x' 表示它的对径点,D_x 表示从 x' 到 x 的直径.我们以后要证明的第一个命题是

(1) 对于任一 $x \in C$,在垂直于 D_x 的全体直线中,有而且只有一条直线 $L(A,x)$ 把 A 按面积分成两半,有而且只有一条直线 $L(B,x)$ 把 B 按面积分成两半.

用 x_A 与 x_B 分别表示 D_x 同 $L(A,x)$ 与 $L(B,x)$ 的交点.在 D_x 上我们有一个以 z 为原点的自然尺度(或坐标系):一点的坐标是它到 z 的有向距离,按照该点与 x 在同侧或异侧,它的坐标就分别是正的或负的.令 g_Ax 与 g_Bx 分别表示 x_A 与 x_B 的坐标.现在,对于每个 $x \in C$,规定

$$hx = g_Ax - g_Bx$$

我们以后要证明的第二个命题是

(2) 函数 $h: C \to \mathbf{R}$ 连续.

h 的一个决定性的性质是：它在圆的任两对径点处的值，绝对值相等而符号相反，即

$$hx' = -hx, \text{对于任何 } x \in C$$

这个式子的证明如下. 注意到 $D_{x'} = D_x$，故 $L(A, x) = L(A, x')$，$L(B, x) = L(B, x')$；因而 $x'_A = x_A$ 与 $x'_B = x_B$. 但是，在 $D_{x'}$ 上，坐标的正向与 D_x 上的相反；从而 $g_A x' = -g_A x$ 与 $g_B x' = -g_B x$，所以

$$hx' = g_A x' - g_B x' = -g_A x + g_B x = -hx$$

现在，由定理 10.1，存在 C 的一点 x 使得 $hx' = hx$. 对于这个 x，$hx' = hx$ 与 $hx' = -hx$ 都成立；因而 $hx = 0$，这蕴涵着 $x_A = x_B$. 于是 $L(A, x) = L(B, x)$ 这同一条直线把 A 与 B 按面积各分成两半.

命题(1)的证明

给定一数 y，D_x 上就有一个以 y 为坐标的点. 令 L_y 表示通过这点并垂直于 D_x 的直线，又令 fy 表示 A 在 L_y 正侧（在 y 增值方向的一侧）的那部分面积. 于是 f 是一个实变数的实值函数，$f: \mathbf{R} \to \mathbf{R}$. 当 y 从 $-r$ 变到 r 时，直线 L_y 把 C 的内部扫过一遍. 设想 L_y 是一条钢针架在横槽 D_x 上，并且与 D_x 成直角. 当钢针的支点沿着横槽 D_x 从 x' 到 x 时，这钢针横扫 C 的内部. 当 $y = -r$ 时，钢针在 x' 处，全部 A 在正的一侧，所以 $f(-r)$ 是 A 的面积. 当 $y = r$ 时，钢针在 x 处，全部 A 在负的一侧，所以 $fr = 0$.

为了证明 f 连续，令 $y, y' \in \mathbf{R}$，其中 $y < y'$. 于是 $fy - fy'$ 是 A 在直线 L_y 与 $L_{y'}$ 之间的那部分面积. 因为这是包含在图 11.2 中用阴影表出的矩形域中，从而 $|fy - fy'| < 2r|y - y'|$. 对应于一个 $\varepsilon > 0$，我们取 $\delta = \varepsilon / 2r$. 于是，当 y' 在 y 的 δ 邻域中时，fy' 就在 fy 的 ε 邻域中. 所以 f 在 y 处连续. 因为对每个 y 都如此，所以 f 连续.

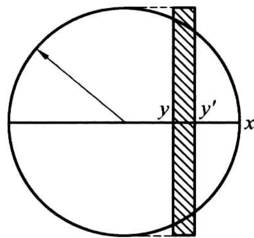


图 11.2

根据主要定理，当 y 从 $-r$ 到 $+r$ 变化时，fy 取遍从 A 的面积开始直到零的诸值. 因此至少有一个 y 值，使 fy 正好是 A 的面积的一半，因而 $L_y = L(A, x)$ 把 A 按面积割成两半. 我们还需要证明只有一个这样的分割. 用反证法：设 L_y 与 $L_{y'}$ 都把 A 分成两半（即 $fy = fy'$），而且 $y \neq y'$，不妨设 $y < y'$. 令 Q 表示介于 L_y 与 $L_{y'}$ 之间的带形区域，它是一个开集，它的补集分成两部分，一份包含 $L_{y'}$ 的正侧而另一份包含 $L_{y'}$ 的负侧. 因为 A 是连通的并且包含这两部分的每一份中的点，A 必包含 Q 的一点，譬如 p. 因为 A 与 Q 都是开集，$A \cap Q$ 是开集，所以它包含 p 的一个邻域. 所以 $A \cap Q$ 有正的面积，因而 $fy > fy'$. 这与 $fy = fy'$ 矛盾，就证明了唯一性. 同样可以证明

$L(B,x)$ 的存在与唯一性.命题(1)证毕.

命题(2)的证明

因为 h 是差 $g_A - g_B$,只要证明 g_A 与 g_B 连续就行了.令 c 为 C 的一点;我们要证明 g_A 在点 c 处连续.又按照上面的记号,令 c_A 为直径 D_0 上的那个点,在该点处垂线 $L(A,c)$ 把 A 切成两半(参看图 11.3).令 x 为 C 的挨近 c 的一点.过 $L(A,c)$ 与 C 的交点 u 与 v,分别作 D_x 的垂线 K 与 K'.直线 $L(A,c)$ 把 C 的内部分成 U 与 V 两部分.介于 K 与 K' 之间的带形窄条把它在 C 内部中的补集分成 U' 与 V' 两部分,使 $U' \subset U$ 与 $V' \subset V$.所以 U' 与 V' 的每一个至多包含 A 的面积的一半.从而垂直于 D_x 的并且把 A 分成两半的直线 $L(A,x)$ 在这窄条中,因而 $L(A,x)$ 与 D_x 的交点 x_A 也在这窄条中.因为过点 c_A 的、以 z 为中心的圆与 D_x 相遇于这窄条的内部,从而

$$| g_A x - g_A c | < w$$

其中 w 是这窄条的宽度.

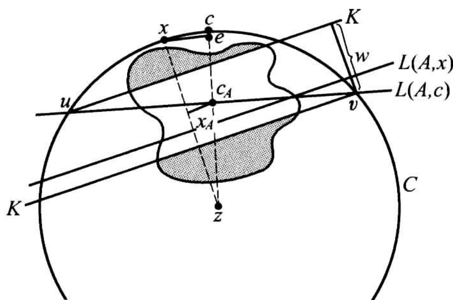


图 11.3

要得到 w 的大小的估计,注意由两个三角形的相似得

$$\frac{w}{d(u,v)} = \frac{d(e,x)}{d(z,x)}$$

其中 e 是 x 到 D_c 的垂足.因为 $r = d(z,x)$,上式是

$$w = \frac{d(u,v)}{r} d(e,x)$$

因为 $d(u,v) \leqslant 2r$,并且 $d(e,x) \leqslant d(c,x)$,得

$$w \leqslant 2d(c,x)$$

所以

$$| g_A x - g_A c | \leqslant 2d(c,x)$$

若 $\varepsilon > 0$,而 $x \in N(c, \frac{\varepsilon}{2})$,从而

$$| g_A x - g_A c | < \varepsilon$$

这就证明了 g_A 连续.同样地 g_B 连续.这完成了命题(2)与定理 11.1 的证明.

关于薄煎饼的第二个问题,要求沿着两条垂直的直线,把薄饼两刀切成四

块相等的部分.

定理 11.2 如果 A 是平面中的有界区域,那么存在着两条垂直的直线,把 A 分成面积相等的四份.

如同先前一样,把 A 放在一个圆 C 的内部.对于每个 $x \in C$,令 L_x 为垂直于 D_x 的分 A 成两半的直线,又令 K_x 为平行于 D_x 的分 A 成两半的直线.这两条线把 A 分成四份,它们的面积按逆时针方向依次用 P_x,Q_x,R_x,S_x 表示(图 11.4).因为 L_x 与 K_x 各把 A 分成两半,我们有

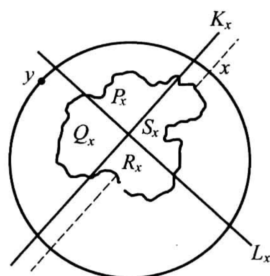


图 11.4

$$P_x + Q_x = R_x + S_x$$

与

$$Q_x + R_x = S_x + P_x$$

这些等式相减,得 $P_x = R_x$ 与 $Q_x = S_x$.如果碰着好运气还有 $P_x = Q_x$,则直线 L_x 与 K_x 就解决了问题.但一般并不是这样;在 P_x 不等于 Q_x 时,设差 $P_x - Q_x = fx$;要问:当 x 绕圆移动时这函数 fx 怎样变化.如果 $y \in C$ 是这样一点,它使 D_y 垂直于 D_x,则显然有 $L_y = K_x$ 与 $K_y = L_x$.从而 $P_y = Q_x$ 与 $Q_y = R_x$.因为 $P_x = R_x$,得

$$fy = P_y - Q_y = Q_x - P_x = -(P_x - Q_x) = -fx$$

所以当 x 移动一段 90° 弧时,函数 f 改变了符号.一旦证明了 f 连续,就会从主要定理得出结论:在每一段 90° 弧的某点处,$fx = 0$.这样的一点提供了所求的切割.

我们只简略地描述一下连续性的证明.因为 f 是两函数的差,只要证明 P_x 连续就行了(Q_x 连续的证明类似).令 $c \in C$ 是我们要证明 P_x 在那里连续的一点,又令 x 为挨近 C 的一点.从一对垂线 L_c,K_c 过渡到类似的一对垂线 L_x,K_x,能用两步完成.首先绕着 L_c 与 K_c 的交点 p,把 L_c,K_c 旋转到分别与 L_x,K_x 平行的一对垂线 L'_c,K'_c.旋转角 α 是 C 上从 c 到 x 的劣弧所对的角.第二步是把 L'_c,K'_c 平移到 L_x,K_x.可以看到,从 P_c 到 P'_c 的改变量不会大于顶点在 P 而角为 α 的 C 的偏心扇形的面积,这面积至多是 $2rd(c,x)$,其中 r 是 C 的半径.介于 L'_c 与 L_x 之间并且在 C 内部的面积 U 至多是 $2ru$,其中 u 是 L'_c 与 L_x 之间的距离.同样,介于 K'_c 与 K_x 之间的面积 V 至多是 $2rv$.从 P'_c 到 P_x 的改变量显然小于

$$U + V \leqslant 2r(u + v)$$

还可以看出,L_c 与 L_x 的交点 q 是在 C 的内部,因为 L_c 与 L_x 分 A 成两半并且 A 是连通的.这说明 $d(p,q) < 2r$;于是,由于相似三角形,$u < 2d(c,x)$.同样,$v < 2d(c,x)$.把这些估计放在一起,就得出

$$| P_x - P_c | < 10rd(c,x)$$

所以,如果给定一数 $\varepsilon > 0$,取 $\delta = \varepsilon/10r$,则对于每一 $x \in N(c,\delta)$,都有 $|P_x - P_c| < \varepsilon$.证毕.

12 多项式的零点

下一个定理是主要定理的代数学上的应用.

定理 12.1 一个实系数的奇次多项式至少有一个实零点.

要领会这定理的含义,让我们观察偶次多项式与奇次多项式的一些特例.首先,若多项式为一次,$fx = ax + b, a \neq 0$,则 $y = ax + b$ 的图形是一条直线,它交 x 轴于 $x = -\dfrac{b}{a}$,所以多项式以 x 的这个值为零点.其次,考虑抛物线 $y = x^2 + 1$,作为二次多项式的一个例子(图 12.1).这曲线完全位于上半个坐标平面内.$x^2 + 1$ 的最小值是 1,因为对于任何实数 x,$x^2 \geqslant 0$;因而多项式没有实零点.同样,$x^6 + 1$ 没有实零点;$x^4 - 2x^2 + 5$ 也没有,因为

$$x^4 - 2x^2 + 5 = (x^2 - 1)^2 + 4$$

绝没有一值小于 4(图 12.2).另一方面,偶次多项式 $x^2 - 4x + 3$ 有实零点 $x = 1$ 与 $x = 3$(图 12.3).

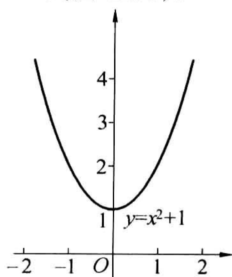


图 12.1

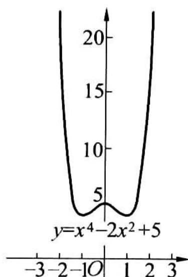


图 12.2

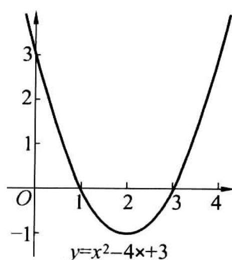


图 12.3

$y = x^3 - x + 5$ 在 -2 与 -1 之间的某处穿过 x 轴.多项式 $x^5 - 2x^3 + x + 4$ 的次数为 5;它在 -1.7 与 -1.6 之间的某处穿过 x 轴.

在我们的例子中,每一个奇次多项式的图形从 $-\infty$ 上升,穿过 x 轴,最后趋向 $+\infty$.偶次多项式的图形从 $+\infty$ 下降,再回升到 $+\infty$,其间可能有一些摆动,在我们的例子中,有些从不穿过 x 轴.定理 12.1 的要点是:奇次多项式决不会这样;每一个实系数的奇次多项式至少有一个实零点.

要证明定理 12.1,只需要考虑形如

$$f(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$$

的多项式.因为如果最高次项的系数不是 1,能用这系数的倒数遍乘多项式的每项而不至于改变多项式的零点.对于 $x \neq 0$,可以把 $f(x)$ 写成

$$x^n\left(1+\frac{a_1}{x}+\cdots+\frac{a_{n-1}}{x^{n-1}}+\frac{a_n}{x^n}\right)$$

或 $f(x) = x^n q(x)$,其中

$$q(x) = 1+\frac{a_1}{x}+\cdots+\frac{a_{n-1}}{x^{n-1}}+\frac{a_n}{x^n}$$

我们证明的方法在于说明奇次多项式 f 在某些 x 处是负的,在另外一些 x 处却是正的,并且是连续的.于是从主要定理就会获得所寻求的结论.

如果 x 是这样一个数,使得

$$\frac{a_1}{x},\frac{a_2}{x^2},\cdots,\frac{a_n}{x^n}$$

的每一项的绝对值小于$\frac{1}{n}$,那么这 n 项的和

$$h(x) = \frac{a_1}{x}+\frac{a_2}{x^2}+\cdots+\frac{a_n}{x^n}$$

的绝对值小于$\frac{n}{n} = 1$;这就是说 $h(x)$ 是介于 -1 与 $+1$ 之间,又因为 $q(x) = 1+h(x)$,$q(x)$ 是正的.为找出使这成立的一个数 x,先看

$$n|a_1|,(n|a_2|)^{\frac{1}{2}},\cdots,(n|a_n|)^{\frac{1}{n}}$$

中的每一个数,选出比它们都大的一个数 b.要说明对于一个使 $|x|\geqslant b$ 的 x,$q(x)$ 是正的,只需用到下列事实:不等式

$$|x|>n|a_1|,|x|>(n|a_2|)^{\frac{1}{2}},\cdots,|x|>(n|a_n|)^{\frac{1}{n}}$$

蕴涵

$$\left|\frac{a_1}{x}\right|<\frac{1}{n},\left|\frac{a_2}{x^2}\right|<\frac{1}{n},\cdots,\left|\frac{a_n}{x^n}\right|<\frac{1}{n}$$

对于使 $|x|\geqslant b$ 的 x 值,多项式的符号就是 x^n 的符号,因为 $f(x) = x^n q(x)$ 而 $q(x)$ 是正的.因为 n 是奇数,x^n 具有 x 的符号.于是 $x = b$ 时多项式是正的,而 $x = -b$ 时多项式是负的.

要应用主要定理来断定 $-b$ 与 b 之间存在一个零点,必须证明多项式是一个连续函数.在第3节中已指出任一常数函数(即零次多项式)与任一恒等函数(即一次多项式)都是连续的.连续函数的积是连续的.从而 $x^2 = x\cdot x$ 连续,$x^3 = x^2\cdot x$ 连续,并且由归纳法,对于每一个 kx^k 都连续.因为 x^k 与常数 a 都连续,同一结果告诉我们任一单项式 ax^k 连续.每一个多项式是它的诸单项式的和,而连续函数的任何和仍连续;所以每个多项式连续.





13 平面的自映射①

从本节开始的目的是证明一对联立方程的解的存在定理. 这个定理在第18节中叙述,它的证明在第26节中完成. 第27节一直到第36节是应用这定理到映射的不动点、向量场的奇点以及多项式的零点等问题上. 为了提出这主要定理,必须发展前面的一维概念的二维类似. 所需的关键概念是平面中一闭曲线围绕不在该曲线上一点的圈数. 我们将首先给出这概念的一个直觉定义连同主要定理的一个直观证明(第17,18节). 在第19 ~ 26节中再使定义精确,证明严密.

记住,前面12节主要定理讨论的是从一线段到一直线的映射 $f:[a,b]\to\mathbf{R}$,还给出点 $y\in\mathbf{R}$ 所应满足的条件,使得在这条件下能断言 y 是在象 $f[a,b]$ 中(即 $fa\leqslant y\leqslant fb$). 本节开始的主要定理将讨论从平面 $P(=\mathbf{R}^2)$ 的一部分 D 到 P 的映射 $f:D\to P$,并给出能断定一点 $y\in P$ 在象 fD 中的条件②.

还应记住,我们看到 $f:[a,b]\to\mathbf{R}$ 的图形这一概念在说明主要定理的意义以及在几何直观地揭露它的真实性方面都非常有用. 在二维空间中我们也可以谈映射 $f:D\to P$ 的图形. 但必须注意它牵涉到这样一个事实:平面 $P=\mathbf{R}^2$ 的一个点由两个实数 (x_1,x_2) 表示,它的 f 象需要另两个,(y_1,y_2). 于是这点以及它的象所组成的这一对点由四个数表示,并且图形上的一个点是四维空间的一点. 这样,f 的图形是 $\mathbf{R}^4$ 中的一个曲面.

这里是我们遇到的第一个困难. 用图形说明我们的定理,需要我们具有能看出四维空间中一个曲面的能力(我们没人有此能力). 所以我们必须采取一种不同的看出映射的方法:如同第2节中所简述的用图表示象以及原象的方法. 在这一节的其余部分中,我们将用同样的方法来讨论较复杂的映射. 我们的意图是加强几何直观,并指明后续各定理的一般性程度.

在第2节中我们讨论了平移、旋转、反射以及相似变换作为平面的自映射. 一个较复杂的映射是:它在一个方向放大长度而在另一方向缩小之. 图13.1说明一个映射 f,它使水平方向的长度扩大一倍而使垂直方向的长度缩小一半. 显然它改变了角与形状. 它把一圆映成一椭圆. 惊奇的是它把任一直线映成直线.

图13.2说明一个剪切变换 $P\to P$. 设想有一个格子栅拦,是用许多水平的与铅垂的长而窄的木条做成的,水平木条与铅垂木条的每一交点处用一个钉子

---

① 以第13节到第37节为二维时的存在定理.

② 在引言中我们说过本定理是讨论一对联立方程 $f(x,y)=a$ 与 $g(x,y)=b$ 的;这形式现转变为目前的形式,因为我们用记号 (x_1,x_2) 替换了 (x,y),用 (y_1,y_2) 替换了 (a,b),用 (f_1,f_2) 替换了 (f,g),并且把数对 (x_1,x_2),(y_1,y_2) 作为 P 中点 x,y 的坐标.

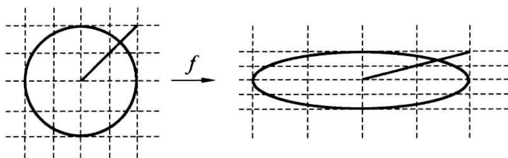


图 13.1

钉着.这样一个结构不是刚体的,而在手指碰着它时,它就能回转,作一个剪切变换.一个剪切也把圆映成椭圆,把直线映成直线.

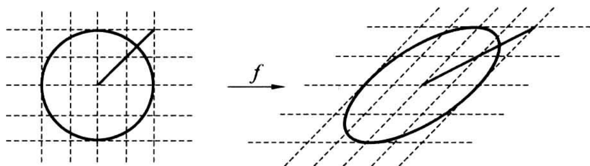


图 13.2

直到现在为止,所考虑的映射都是一对一的.我们也要考虑不是一对一的映射.图 13.3 表明沿着 P 的一条直线把 P 简单地折叠起来.沿着折叠线,这映射是一对一的,但在折叠线上部的每一点是平面的两个不同点的象.

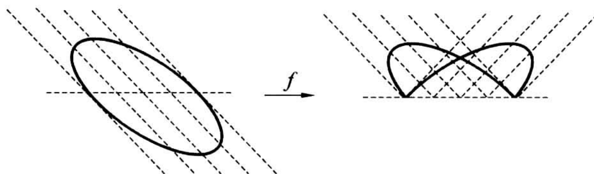


图 13.3

图 13.4 说明一平面盖满它自身两次的一种映射.中心点 z 映成它自身.射线 L 的每一点也固定.每一条以 z 为起点的射线刚体地映成一条以 z 为起点的射线,但后者与 L 的夹角两倍于前者与 L 的夹角.考虑一条以常角速度绕 z 旋转的射线;它的象是用两倍于这速度绕 z 旋转的射线.当第一条射线绕过半转时,它的象完成一个整圈.把 z 除外,这个映射是2到1的.每一个以 z 为中心的圆围绕自身两次.用一个整数 n 乘角速度,就得出一个类似的映射.它是 n 到 1 的,z 除外.

还有更复杂的映射:把 P 缠绕 P 无限多次.图 13.5 就表示出这么一个映射.水平线 L 映成单独一个点 z,而每条铅垂线刚体地映射成过 z 的一条直线.当铅垂线以常速度沿水平方向移动时,它的象以常角速度绕着 z 旋转.此图只显示出 180° 的旋转.这个映射是 ∞ 到 1 的.z 的原象是一整条直线 L.任何其他点的原象由两排孤立的点组成,一排在 L 之上而另一排在 L 之下,一排中每对相邻的点都等距.

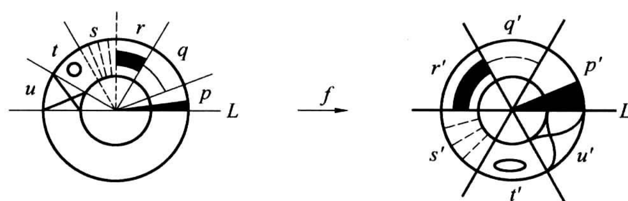


图 13.4

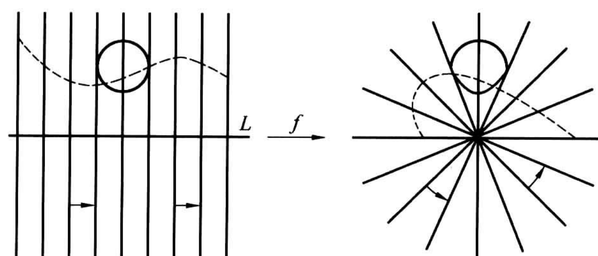


图 13.5

能够用几句话明确地描述的映射,通常都太简单,远不能说明一般情形下的映射的复杂性.图 13.6 说明一个较复杂的映射.我们不详细地描述它,只说:它把一族同心圆映射成一族 8 字形;P 的象正好是 P 在两条射线所围区域中的那一部分.直观地说,P 的象可以看做是沿着一条直线抖出来的一串同心圆,其中的每个圆还都同时扭弯成 8 字形.

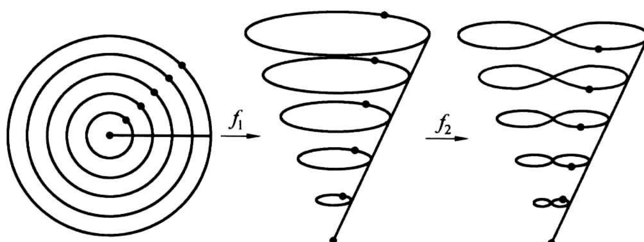


图 13.6

14 圆 片

在一维空间诸定理的叙述与证明中,区间是一个重要角色.对于二维空间的定理,圆片将起类似的作用.我们规定:平面中一圆 C 与它的内部的并集叫做一个圆片 D.圆 C 叫做 D 的边界.圆片 D 由它的中心 z 与它的半径 r 确定.点在圆片中如果它与 z 的距离小于或等于半径;也就是,$x \in D$ 指的是 $d(x, z) \leqslant r$.

我们已注意到,任两区间是相似的,所以是拓扑等价的;任两圆片 D 与 D' 也如此;如果 D 与 D' 的中心不是同一点,则可以把 D' 平移到与 D 有同一中心的 D'' 处.于是,在 z 处一个适当的放大或缩小将把 D'' 映成 D.

在一维时,直线的任一子集若与一区间拓扑等价,则这子集本身是一个区间,因为它必须是紧致的、连通的.但在平面中就有许多很不相同的子集与一个圆片拓扑等价.例如在一个剪切变换下,一圆片能映成一椭圆与它的内部.图14.1说明一圆片映成一简单闭曲线与它的内部的一个拓扑映射.中心 z 映成 z',而 D 的每条径向线段 zy 由一相似变换(放大或缩小)映成平行线段 $z'y'$.对于任何凸多边形,例如三角形或矩形,也可以用同样的办法.

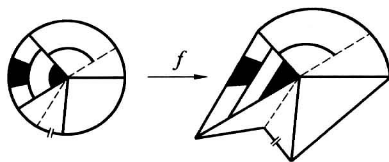


图 14.1

在二维时的存在定理中,能把圆片这个名称解释为,与圆片拓扑等价的这些图形中的任一个.但因为圆片的对称性以及易于描述,我们宁愿选择它来代表其他拓扑等价的图形.

圆片 D 当然是有界集,因为它处于任一个以它的中心为中心而半径较大的圆之内.它也是一个闭集,因为它的补集的每个点有一个邻域不与 D 相遇;事实上,若 y 不在 D 内,则 $d(y,z)-r$ 大于 D 的半径,从而半径为 $r'=d(y,z)-r$ 的 y 的圆邻域就不包含 D 的点.因为 D 是闭的而且有界,从而 D 是一紧致集.于是,在任一映射 $f:D\to P$ 作用下,象 fD 是紧致的,所以是闭而有界的.

对于 D 的任意两点,连接它们的线段也在 D 中.所以 D 是一个连通集.于是,对于任一映射 $f:D\to P$,fD 是连通的.

15 主要定理提法的初步尝试

二维时的主要存在定理类似于一维时的主要定理.它说:如果 $f:D\to P$ 是从圆片到平面的映射,则对于满足一定条件的每一点 $y\in P$,方程 $fx=y$ 有一个解 $x\in D$.这个条件的提法颇为复杂.我们将先提出几个简单而貌似合理的条件,然后说明它们为何不够恰当;经过这么几个阶段来获得明确提法.

在一维的定理中,D 是闭区间$[a,b]$,y 所满足的条件是它介于 fa 与 fb 之间.这里 a 与 b 为区间端点,并且把区间与直线的其余部分分开.在圆片的情形,D 的最外边的点是边界圆 C 的点,而 C 把 D 与平面的其余部分分开.于是要提出的条件可以说成 y 以某种方式与 fC 有关.说"y 介于 fC 间",这明显是胡扯.

如果把一维时的条件改说成:要求 y 被 fa 与 fb 包围着,这跟"介于 fa 与 fb 之间"有同样意义,而二维时的类似语"y 被 fC 包围着"却有了直观的意义.让我们试着确切地提出这个说法.作为尝试的第一步,考虑"y 是以 fC 为边界的圆片的一点".这不恰当,因为对于许多映射 f 来说,fC 不是一个圆.它很容易是一椭圆或一矩形.作为第二步尝试,考虑"y 是以 fC 为边界的区域的一点".这比较好些,但排斥了 fC 是8字形这种可能情形.作为第三步尝试,试用"y 是某个有界区域的一点,这区域的边界包含在 fC 中".在我们检查一下图 15.1 所示的一个 $f: D \to P$ 的例子之前,这说法似乎是所需要的.映射 $f: D \to P$ 最好用从左到右所画出的几步来描述.首先,把 D 拉成一长的窄条 E.其次,把 E 弯成一弯曲形状的 F,它像加厚了的一个圆的四分之三.最后,把 F 继续往下弯,到两端重迭在最后的图形 fD 中为止.标出的点 y 不在 fD 中,可是它属于一个有界区域,其边界在 fC 中.

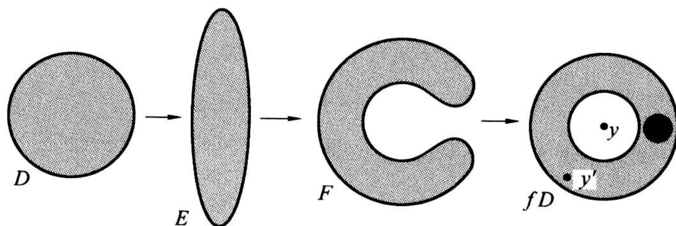


图 15.1

最后这个例子清楚地显示出我们必须克服的困难.fD 中标以 y' 的点与 fC 有什么关系,而点 y 却没有这种关系?答案的提法用到我们将要发展的一个新概念:曲线围绕一点的围绕数.我们将看到 $f \mid C$ 围绕点 y' 的圈数不是零,而它围绕点 y 的圈数是零.这就是为什么 $fx = y'$ 有一个解 $x \in D$,但 $fx = y$ 却无解.

16　曲线与闭曲线

在这以前,"曲线"这个词指的是连续函数 $f:[a,b] \to \mathbf{R}$ 的图形.现在我们需要在下列比较广的意义下用这个词.平面中一条曲线定义为从实数的某个区间 $[a,b]$ 到平面 P 的一个映射 $\varphi:[a,b] \to P$.每一数 $t \in [a,b]$ 可以看成时间的一瞬间,对应点 $\varphi t \in P$ 可以看成一动点在时刻 t 的位置.于是一条曲线可以当做一动点走出来的道路.特别是任何一条曲线有一个定向,意思是选定沿曲线从 φa 到 φb 的方向为正向.这是随着时间增加的运动方向.画曲线的图形时,用箭头表示定向,如图 16.1 所示.注意,我们容许一条曲线自交;即动点能在几个不同时刻通过同一点.再者,动点也可以在一段时间内保持静止.例如,把整个区间 $[a,b]$ 映成单独一点的常值函数就是我们意义上的一条曲线.如果一条

曲线的起点和终点是同一点，即 $\varphi a = \varphi b$，它就叫做闭曲线.



图 16.1

在 P 中，从一点 A 到一点 B 的线段 L 可以表示为一曲线. 记住，任两线段总是相似的. 所以，如果 $\varphi:[a,b]\to L$ 是相似变换，使 $\varphi a = A$ 与 $\varphi b = B$，则 φ 定义一条曲线，它的象是 L. 这例子中的动点具有常速度.

连续函数 $f:[a,b]\to \mathbf{R}$ 的图形是一条曲线. 我们只要对于每个 $t\in[a,b]$，令 φt 是坐标为 (t,ft) 的点即可. 这种类型的曲线既不自交，也不是闭的，因为 $t_1\neq t_2$ 蕴涵 φt_1 与 φt_2 这两点有不同的横坐标.

任一圆 C 按下述标准方式可看做是一条闭曲线. 令 z 表示 C 的中心并令 L_0 表示从 z 出发的一固定射线（包含起点的半直线）. 对于 $t\in[0,1]$，定义 φt 如下：$\varphi 0$ 是 L_0 与 C 的交点，φt 是 C 的这样一点，使连接 z 到 φt 的线段与 L_0 在 z 处的交角是 $360t°$. 例如 $\varphi(\frac{1}{4})$ 是在 $90°$ 处的点（绕圆 $\frac{1}{4}$ 圈处）.（图 16.2）因为整个圆有 $360°$，我们有 $\varphi 1 = \varphi 0$. 在这个情况下，动点也有一个常速度.

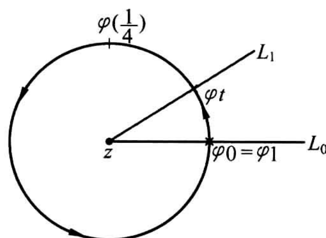


图 16.2

矩形的边界同样可以看做是一条闭曲线. 取一区间 $[a,e]$，用 $a<b<c<d<e$ 的数 b,c,d 把这区间 $[a,e]$ 分成四个子区间. 命矩形的四个顶点按照同样顺序为 A,B,C,D. 如同上例，能定义 φ 使得它把区间 $[a,b]$，$[b,c]$，$[c,d]$ 与 $[d,e]$ 分别映成线段 AB,BC,CD 与 DA. 曲线是闭的，因为 $\varphi a = A$ 与 $\varphi e = A$.

我们用图做的说明，可能会使人们倾向于把一曲线 φ 看做不过是象集 $\varphi[a,b]$；这却是误解. 故必须强调指出曲线是映射 φ. 例如，把圆看做是闭曲线，就有无穷多的标准表示法，L_0 的每一选择就有一个表示.

17　围绕数的直观定义

令 $\varphi:[a,b]\to P$ 为一闭曲线，令 y 为平面 P 的不在曲线上的一点，对于区间中的每个 t，令 L_t 表示从点 y 开始并通过 φt 的射线. 当 t 从 a 变到 b 时，点 φt 描出这曲线，而射线 L_t 绕着它的固定端点 y 旋转. 因为曲线是闭的，L_t 最后回到它的初始位置 $L_b = L_a$. 所以这射线在运动的全过程中，围绕着 y 旋转了若干个整圈. 整圈的个数叫做闭曲线 φ 在点 y 处（或围绕 y）的围绕数，并用简单记号 $W(\varphi,y)$ 来表示这个数. 通常约定，逆时针方向旋转时这个数是正的，顺时针方向旋转时是负的.

在图 17.1 中,把圆看做 16 节所述的一条闭曲线,它围绕它的中心一次.它也围绕圆内任一点一次.但是它围绕圆外任一点的围绕数是零.设 y 是圆外任一点,以 y 为起点的射线中恰有两条与这圆相切,它们小于 180° 的夹角包含这个圆.当 φt 描出这个圆时,固定在 y 处的射线 L_t 扫过平面夹在两切线之间的部分.L_t 先向一个方向扫去;达到一边界(切于圆的射线之一)后就逆转方向,结果是未旋转完一次就回到它的初始位置(图 17.2).

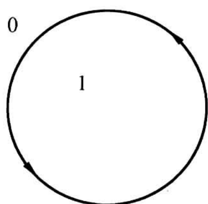


图 17.1

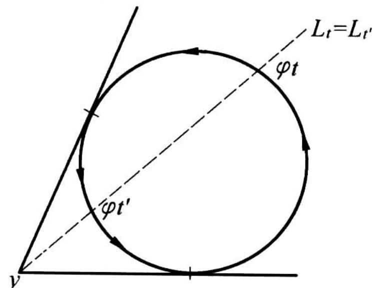


图 17.2

在图 17.3 中,闭曲线是沿顺时针方向描出一次的一个椭圆.对于它内部的任一点,围绕数是 - 1,而对于它外部任一点,围绕数是零.

图 17.4 是一条 8 字形的闭曲线.它在一个有界区域内部的每一点处的围绕数都是 1,而在另一个有界区域内部的每一点处的围绕数却是 - 1.当然在无界区域内的每一点处的围绕数都是零.

图 17.5 表示在 15 节中所讨论的例子.在此例中,存在一有界区域,在其内的点处,曲线的围绕数是零.对于另两个有界区域,围绕数是 1 与 2.注意,在同一个连通集的两点处,围绕数总是同一个数.

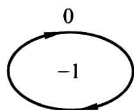


图 17.3

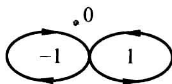


图 17.4

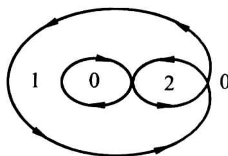


图 17.5

图 17.6 表示常值函数所给出的常曲线(全部只单独的一点).它围绕每个其他的点零次.

图 17.7 指出可能性是无止境的.

现在令 C 为一圆,令 $f: C \to P$ 为从 C 到平面的一映射,令 $\varphi:[0,1] \to C$ 为把 C 看做 16 节所述闭曲线时的标准表示法.于是复合函数 $f\varphi:[0,1] \to P$ 仍是

一闭曲线,因为 $\varphi 0 = \varphi 1$ 蕴涵 $f\varphi 0 = f\varphi 1$.设 y 是 P 中不在 fC 上的任一点,这曲线在 y 处有一围绕数.这叫做 f 在 y 处的围绕数,并用 $W(f, y)$ 表示.

O•

图 17.6

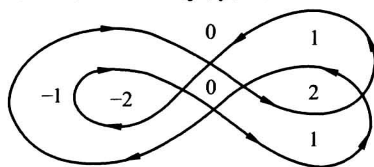


图 17.7

18 主要定理的陈述

利用围绕数这个概念,现在就能确切提出二维时的主要定理.

定理 18.1 设 $f: D \to P$ 是从一圆片到平面的映射,C 是 D 的边界圆,并设 y 是平面中不在 fC 上的一点.如果 $f \mid C$ 在点 y 处的围绕数不是零,则 $y \in fD$;也就是说,存在一点 $x \in D$,使得 $fx = y$.

下面是一个简短的直观证明.设 C 的半径是 r,中心是 z.对于每一个 s,$0 \leqslant s \leqslant r$,用 C_s 表示以 s 为半径的,C 的同心圆;于是 $C_r = C$,C_0 是中心 z.设 y' 是平面上不在 fD 中的一点.则对于$[0, r]$内的每一个 s,y' 不在 fC_s 上,因为 C_s 是在 D 中,因而对于$[0, r]$内的每一个 s,围绕数 $W(f \mid C_s, y')$ 有定义.把它简写为 $W(s)$.考虑当 s 从 r 下降到 0 时的全体闭曲线 $f \mid C_s$.这族闭曲线从 $f \mid C$ 开始,最终缩成常数曲线 $f \mid C_0$,即缩成点 fz.因为 $f \mid C_s$ 随着 s 的不断减小而逐渐变化,从而 $W(s)$ 是 $s \in [0, r]$ 的一个连续函数.围绕数 $W(s)$ 怎样变化呢?答案是:根本不变,因为 W 是 s 的连续函数,而 $W(s)$ 的每个值必是一整数;不可能从一个整数值跳到另一个整数值而不经过其间的非整数值.于是对于所有的 s,$W(s)$ 有同一个值;特别是 $W(r) = W(0)$.但 $W(0) = 0$,因为 $f \mid C_0 = fz$ 是常数函数所给出的闭曲线.所以对于不在 fD 中的每个点 y',$f \mid C_r$ 在 y' 处的围绕数是零.从而 $W(f \mid C_r, y) \neq 0$ 蕴含着 y 是在 fD 中;而说 y 在 fD 中就是说存在一个 $x \in D$ 使 $fx = y$.

在图 18.1 的说明中,要想看出上述论点是怎样引进的,只要看一看 s 递减的几个阶段中的闭曲线 $f \mid C_s$.$f \mid C_r$ 有两叶(看图 15.1 中的 fD,图 18.1 中的粗曲线 $f \mid C_r$);一旦这两叶分开(例如,图中画出的第三条闭曲线),y 明显地在这个闭曲线之外,从而在 y 处的围绕数为零.注意,这与图 17.5 所示结果一致.

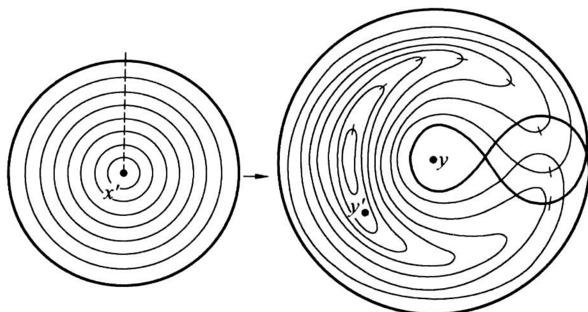


图 18.1

19 什么时候论点算不了证明?

绝大多数人在看完并且理解了前两节中的论点之后,都认为论点有说服力,定理已是真理,无需再加什么来获得一个完整的、合乎逻辑的证明.但是细心的读者会发现论证上的缺陷.主要的缺陷发生在第 17 节;没给出围绕数的一个精确定义.在那一节里,当 t 从 a 变到 b 时,射线 L_t 绕着它的起点 y 究竟旋转多少个整圈,留待直观去决定;这是假设了我们的眼睛能跟着旋转的射线并把它的运动总数加成单一个旋转圈数.大家知道,眼力在这种场合并非完全可靠;例如,足够快地映出一串静止画面时,会使我们产生一个错觉,认为是在看连续的动作.

好在数学概念与推理并不依赖于我们看见动作的能力.我们必须处理的是静止的情形.我们有一条闭曲线 φ,有不在这曲线上的一点 y,我们想要赋予 y 与 φ 一个整数叫做围绕数,使得这名称符合于我们的直观看法.这将在以下 7 节中去做.如果读者要尽早看一些新概念和新应用,不准备停留在已勾画出来的概念的细致发展,他应该跳到第 27 节,而删去以下的 7 节.

在我们集中注意力到 $W(\varphi,y)$ 定义的细节之前,我们要指出,为完成主要定理的陈述与证明,只需要:

(1) 精确地定义 $W(\varphi,y)$,

(2) 在第 18 节的直观证明中所用的 φ 的那种变化下,证明它是连续的,以及

(3) 证明只要 φ 是常值闭曲线,就有 $W(\varphi,y)=0$.

如果我们采取的定义是:$W(\varphi,y)=0$ 对于全体 φ 与 y,那么这个定义自然满足(1),(2) 与(3) 三个要求,所以主要定理的证明对于这个 W 有效;但是定理的结论什么都没说,因为不存在点 y 使 $W(f\mid C,y)\neq 0$.于是,为着使得我们的努力不是白费,还要增加下述的要求:

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

(4) 对于某些曲线 φ 与点 γ, $W(\varphi,\gamma)$ 应该不是零;特别是,它应该与第 17 节中直观地定义的围绕数一致.

20 一曲线所扫过的角

为了提出围绕数的一个良好定义,首先考虑"一曲线 $\varphi:[a,b]\to P$ 在一点 γ 处所扫过的角" 这个更一般的概念. 我们将分两步来定义这个角的大小 $A(\varphi,\gamma)$: 先对特殊类型的曲线,后对任意的曲线.(除非另有声明,我们都用度数来量角而略去度的记号°) 于是我们将看出,一条闭曲线所扫过的角,其大小必是 360 的倍数,而用这个倍数来定义围绕数 $W(\varphi,\gamma)$

$$W(\varphi,\gamma)=\frac{A(\varphi,\gamma)}{360}$$

有一种特殊类型的曲线,即,由所谓短曲线组成的曲线,我们能很容易而明确地定义它们在曲线外一点 γ 处的$A(\varphi,\gamma)$. 称 φ 关于不在φ 上的点γ 是短的,如果存在一条射线 L,它以 γ 为起点并且不交 φ. 例如,一个常函数所给出的曲线即点曲线 $\varphi t=z\neq\gamma$ 对于所有$t\in[a,b]$,它就是关于 γ 的短曲线. 图 20.1 给出一个稍为复杂的短曲线的例子. 当 t 从a 变化到 b 时,φt 沿曲线从 φa 变化到φb,而从 γ 发出的到曲线上点φt 的射线从L_a 旋转到 L_b. 令 $\sphericalangle L_aL_b$ 表示这个旋转角;因 φ 是短的,这个角内不会包含射线 L(从 γ 开始,不与 φ 交). 定义 $A(\varphi,\gamma)$ 为 $\sphericalangle L_aL_b$ 的度数;逆时针方向的角是正号而顺时针方向的角是负号.

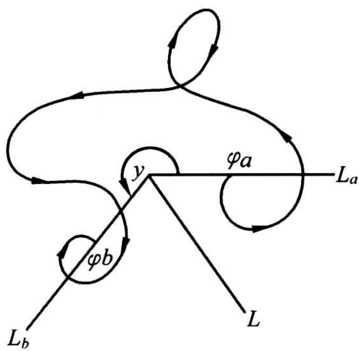


图 20.1

设想我们有一量角器,它的形状是一个整圆周 C 被分成 360 相等的弧,圆周上从一个零点开始依逆时针方向刻出从 0 到 359 的点. 我们这样放置 C,使它的中心在点 γ 处,而沿射线 L 是零度(即 C 上的零点是C 与L 的交点). 令 p_a, p_b 依次表示 C 与射线L_a, L_b 的交点,又令 x_a, x_b 表示它们在量角器上的相应读数,则 $\sphericalangle L_aL_b$ 的度数能由下式算出

$$A(\varphi,\gamma)=x_b-x_a$$

注意:这个差 x_b-x_a 不依赖于量角器起始射线的位置,只要这个射线不被包含在所要测量的角 A 之内(图 20.2). 这一事实可证明如下. 假定把量角器旋转一个角度 α 使得它现在的零度在射线 L' 处. 这新的量角器读数是

$$x'_a = x_a - \alpha \quad 与 \quad x'_b = x_b - \alpha$$

结果 $x'_b - x'_a = x_b - x_a = A(\varphi, y)$ 只要 L, L' 不与作为曲线径向射影象的 C 的这段弧 $p_a p_b$ 相交. 容易看出在这些条件下,量角器的读数总是介于 0 与 360 之间的数,而差 $x_b - x_a$ 总是在 -360 与 360 之间. 在图 20.1 中,这个差是正的,并且大约是 230.(如果曲线的定向反过来,即如果 φa 与 φb 互换,则 $A(\varphi, y)$ 大约是 -230)于是公式

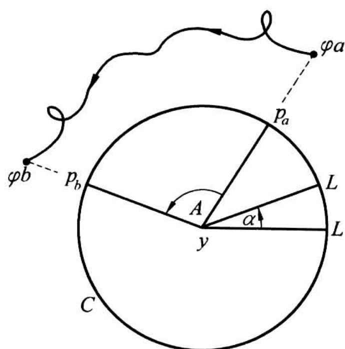


图 20.2

$$A(\varphi, y) = x_b - x_a$$

唯一地给出由短曲线 φ 在点 y 处所扫过的角的度数. 此外,"短" 曲线的定义保证了存在一条以 y 为起点的不与 φ 相交的射线 L,L 可以作为我们计算 $A(\varphi, y)$ 的量角器的起点;如果有许多这样的射线,选择哪一条作为我们量角器的起始射线都无所谓.

留心我们的定义怎样提供正向与负向运动的相互抵消. 例如在图 20.1 中,射线 L_t 开始旋转一个负 30° 角,但又立即转回到初始位置,因而正负相销了. 类似地,当点经过曲线顶上的圈时,L_t 旋转角的总和为零.

如果 φ 是一个点曲线,把$[a, b]$映成一个单独的点,则 φ 是一短曲线并且 $A(\varphi, y) = 0$;因为 $L_a = L_b$,因而

$$x_a = x_b$$

函数 $A(\varphi, y)$ 有一最重要的性质,叫做关于 φ 的可加性. 令 $a < b < c$ 为实数,又令 $\varphi:[a, c] \to P$ 为关于 y 的短曲线. 令 φ_1, φ_2 为 φ 的两部分,其中 φ_1 是 φ 在$[a, b]$上的限制而 φ_2 是 φ 在$[b, c]$上的限制. 我们可把 φ 设想为 φ_1 与 φ_2 的并集. 显然 φ_1, φ_2 也都是关于 y 的短曲线. 选择一条以 y 为起点而不与 φ 相交的射线 L,对于射线 L_a, L_b, L_c 得到它们在量角器上的读数依次为 x_a, x_b, x_c. 因

$$x_c - x_a = (x_b - x_a) + (x_c - x_b)$$

故

$$A(\varphi, y) = A(\varphi_1, y) + A(\varphi_2, y)$$

21 把一曲线划分成短曲线

如果 $\varphi:[a, b] \to P$ 是一曲线,则把$[a, b]$分成子区间,依次取 φ 在每一子区间上的限制,就可把曲线 φ 分成若干段,这样,φ 就可分解成若干较短曲线的并集. 若 φ 不是关于点 y 的短曲线,则会出现这样的情况:在这样一个分解中,φ

的每一个段是关于 y 的短曲线. 于是,把各段曲线在点 y 处所张角的度数加起来,就可得到 $A(\varphi,y)$ 的值.

说得更精确些,把曲线 $\varphi:[a,b]\to P$ 分成一些曲线的并集的分解叫做 φ 的一个划分 $\wp$. 组成划分 $\wp$ 的首先是数的一个增序列,从 a 开始到 b 为止

$$a = t_0 < t_1 < \cdots < t_{m-1} < t_m = b$$

其次是曲线序列 $\varphi_1,\varphi_2,\cdots,\varphi_m$,其中 φ_i 为 φ 在区间 $[t_{i-1},t_i]$ $(i=1,2,\cdots,m)$ 上的限制. 如果每一段 φ_i 是关于 y 的短曲线,我们就说这划分在 φ 外一点 y 处充分地细. 在这种情形下,$A(\varphi_i,y)$ 的每一个有定义,它们的和用 $A(\wp,y)$ 表示

$$A(\wp,y) = \sum_{i=1}^{m} A(\varphi_i,y) = A(\varphi_1,y) + A(\varphi_2,y) + \cdots + A(\varphi_m,y) \qquad (21.1)$$

我们要证明两个命题:

(1) 如果 φ 是任一曲线而 y 是不在 φ 上的一点,则存在 φ 在 y 处充分地细的一个划分.

(2) 如果 $\wp$ 与 $\wp'$ 是 φ 在 y 处充分地细的任两划分,则

$$A(\wp,y) = A(\wp',y)$$

一旦我们证明了这些事实,就能对于任一曲线 φ 定义 $A(\varphi,y)$ 如下:

定义 如果 φ 是平面中的一曲线而 y 是不在 φ 上的一点,对于 φ 的所有充分细的划分 $\wp$,$A(\wp,y)$ 的公共值是 φ 在 y 处所扫过的角的度量. 这个角用 $A(\varphi,y)$ 表示,并可从公式(21.1)算出,右端和式中的每一项由第 20 节的方法算出.

第一个命题告诉我们,能找到一个划分 $\wp$,使 $A(\wp,y)$ 有定义. 第二个命题告诉我们,这样得出的数 $A(\wp,y)$ 不依赖于我们选取的划分,因而它只依赖于 φ 与 y.

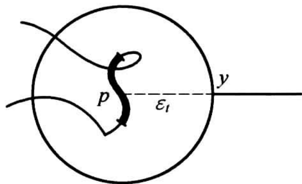
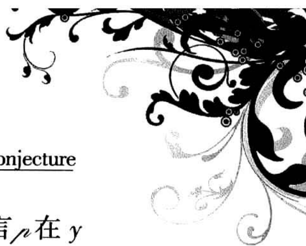


图 21.1

命题(1)的证明 在曲线 φ 的任意点 $p=\varphi t$,以 p 为圆心以 $d(p,y)$ 为半径的圆通过 y(图 21.1). 曲线 φ 在这圆内部的任一段,是关于 y 的短曲线,因为它与从 y 发出的并是线段 py 在圆外的延伸的射线不相交. 应用 φ 在 t 处的连续性,取 $\varepsilon_t = d(p,y)$. 它提供一个数 $\delta_t>0$,使得对于每一个 $t'\in N(t,\delta_t)$,得到 $\varphi_{t'}\in N(p,\varepsilon_t)$. 因此对于每个区间 $I'\subset N(t,\delta_t)$,曲线 $\varphi \mid I'$ 是关于 y 的短曲线.

因 $[a,b]$ 是紧致的,并且邻域 $\{N(t,\delta_t)\}$ 覆盖着 $[a,b]$,就存在这些邻域的有限个,比如说,$N_1,N_2,\cdots,N_k$,覆盖着 $[a,b]$. 令 S 为开区间 $N_1,N_2,\cdots,N_k$ 的所有端点的闭集. 对于 $s,t\in S$ 而 $s\neq t$,命 δ 为所有距离 $d(s,t)$ 中的最短的一





个的一半. 令$\mathscr{P}$为$[a,b]$的任意一个划分,其区间长度至多为δ. 我们断言$\mathscr{P}$在y处是充分地细的. 为了证明这事实,只要指出$\mathscr{P}$的每个子区间I'被包含在N_1, $N_2,\cdots,N_k$的某一个中. 因为I'的长度至多是δ, I'或不包含S的点,或包含S的一个点,设为C. 在第一种情形,取任一与I'相交的N_i(I'是被$N_1,\cdots,N_k$所覆盖);于是$I'\subset N_i$,因为区间I'不包含开区间N_i的端点. 在第二种情形,取任一包含c的N_i;则仍有$I'\subset N_i$,因为I'不包含N_i的端点. 这就完成了命题(1)的证明.

命题(2)的证明 设$\mathscr{P}$是φ的一个充分细的划分. 设在$\mathscr{P}$的一子区间I_k内引进一个新点把I_k分成两个子区间I', I'',因而从$\mathscr{P}$获得一个划分$\mathscr{P}'$. 则在第20节所证明的可加性告知我们

$$A(\varphi \mid I', y) + A(\varphi \mid I'', y) = A(\varphi \mid I_k, y)$$

两端同时加上所有$j \neq k$的各项$A(\varphi \mid I_j, y)$,就有结果

$$A(\mathscr{P}, y) = A(\mathscr{P}', y)$$

把划分$\mathscr{P}$的点$\{t_0, t_1, \cdots, t_m\}$叫做$\mathscr{P}$的顶点. 如果$\mathscr{P}, \mathscr{P}'$是$\varphi$的两个划分,使得$\mathscr{P}'$的顶点包含$P$的顶点,则$\mathscr{P}'$叫做$\mathscr{P}$的加细,并用$\mathscr{P}' < \mathscr{P}$来表示. $\mathscr{P}'$的每个子区间必在$\mathscr{P}$的某个子区间之内. 所以如果$\mathscr{P}$充分地细,所有$\mathscr{P}$的加细也一样.

如果$\mathscr{P}' < \mathscr{P}$,可取$\mathscr{P}'$的那些不是$\mathscr{P}$的顶点的顶点,每次只引进它们中的一个. 这样,我们就得到一序列的加细$\mathscr{P} = \mathscr{P}_1 > \mathscr{P}_2 > \cdots > \mathscr{P}_s = \mathscr{P}'$. 再假定$\mathscr{P}$是充分细的. 于是这证明的头一段的结果给出

$$A(\mathscr{P}, y) = A(\mathscr{P}_1, y) = A(\mathscr{P}_2, y) = \cdots = A(\mathscr{P}_s, y) = A(\mathscr{P}', y)$$

所以$\mathscr{P}' < \mathscr{P}$与$\mathscr{P}$充分地细蕴涵$A(\mathscr{P}, y) = A(\mathscr{P}', y)$.

现在令$\mathscr{P}_1, \mathscr{P}_2$为任两充分细的划分. $\mathscr{P}_1$与$\mathscr{P}_2$的顶点的并集是另一个划分的顶点集,我们把这另一个划分叫做$\mathscr{P}_3$. 显然$\mathscr{P}_3 < \mathscr{P}_1$与$\mathscr{P}_3 < \mathscr{P}_2$. 前一段的结果提供

$$A(\mathscr{P}_1, y) = A(\mathscr{P}_3, y)$$

与

$$A(\mathscr{P}_2, y) = A(\mathscr{P}_3, y)$$

因而

$$A(\mathscr{P}_1, y) = A(\mathscr{P}_2, y)$$

证明完成.

22 围绕数 $W(\varphi, y)$

一旦找到了φ在y处的一个充分细的划分$\mathscr{P} = \{\varphi_0, \varphi_1, \cdots, \varphi_m\}$,只要按例行的办法,使用量角器来得到各$A(\varphi_i, y)$;并注意应有的正负号而相加其结果就得出$A(\varphi, y)$. 在求$A(\varphi_i, y)$时,我们总把不与$\varphi_i$相交的射线$L_i$放在刻度为零的位置上,然后取$\varphi_i$的两端点的读数的差. 于是对于每一个短曲线,需要把

量角器重新定位.现在我们将说明如何能用量角器的单独一个位置来得到所有的读数,以及如何能把计算方法大为缩短.

跟前面一样,令 C 为以 y 为中心而半径为1的圆.对于各 $i=0,1,\cdots,m$,令 p_i 表示 C 与以 y 为起点并且过曲线 φ 上的点 φt_i 的射线的交点.对各短曲线 φ_i,令 L_i 为以 y 为起点并且不与 φ_i 相交的射线.圆上两点 p_{i-1} 与 p_i 决定两段弧;令 $p_{i-1}p_i$ 表示这两弧中不与 L_i 相交的那一段.按照第20节所说明的,$A(\varphi_i,y)$ 是这弧的角度量,附有适当的符号.弧的定向是由 p_{i-1} 到 p_i.如果这定向系逆时针方向,$A(\varphi_i,y)$ 的符号为正,否则为负.

为了方便,假定量角器的半径为1.把量角器的中心放在点 y 处,并通过旋转把量角器的零放在 C 的一点 q 处,q 与 $p_0,p_1,\cdots,p_m$ 都不同.现保持量角器固定,且令 $x_0,x_1,\cdots,x_m$ 依次表示点 $p_0,p_1,\cdots,p_m$ 的直接读数,以度计.各 x_i 都在0与360之间.现在我们能叙述 $A(\varphi,y)$ 的简化公式.

定理22.1 令 r 为含 q 的正向弧 $p_{i-1}p_i$ 的个数,又令 s 为含 q 的负向弧 $p_{i-1}p_i$ 的个数,则

$$A(\varphi,y)=x_m-x_0+(r-s)360$$

图22.1给出定理的说明.如果画出射线 yq,就看出它遇曲线 φ 三次.在各交点处,φ 的定向,因而弧 $p_{i-1}p_i$ 的定向,都能被决定.先为负,而后再为负,最后为正;故 $r=1$ 而 $s=2$.量角器的读数以 q 为零,对 x_m 为65,对 x_0 为195.故 $A(\varphi,y)=65-195+(1-2)360=-490$.

这定理的一个推论是:虽然 x_m,x_0,r 与 s 确实依赖于零点 q 的选择,但

$$x_m-x_0+(r-s)360$$

这数并不依赖于零点 q 的选择.

为着证明这定理,回忆

$$A(\varphi,y)=\sum_{i=1}^{m}A(\varphi_i,y)$$

各 $A(\varphi_i,y)$ 是弧 $p_{i-1}p_i$ 的角度量,我们将用读数 x_{i-1},x_i 把它表示出来.

首先考虑一个 i,使得弧 $p_{i-1}p_i$ 如图22.2所示,不包含 q.按照一短曲线所扫过的角的定义,$A(\varphi,y)=x_i-x_{i-1}$.注意,即使当弧的定向为负时,这仍成立,因那时 $x_{i-1}>x_i$,故 x_i-x_{i-1} 为负.

其次考虑一个 i,使弧 $p_{i-1}p_i$ 为正向并含有 q,如图22.3所示.由弧 $p_{i-1}q$ 与弧 qp_i 所决定的角相加,得

$$A(\varphi_i,y)=360-x_{i-1}+x_i=x_i-x_{i-1}+360$$



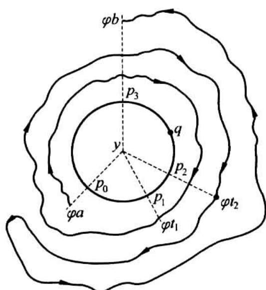


图 22.1

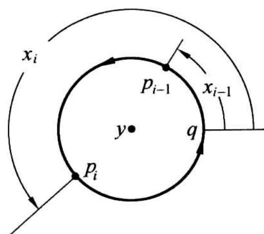


图 22.2

最后考虑一个 i,使弧 $p_{i-1}p_i$ 为负向并含有 q,如图 22.4 所示. 由弧 $p_{i-1}q$ 与弧 qp_i 所决定的角相加,得

$$A(\varphi_i, y) = -x_{i-1} - (360 - x_i) = x_i - x_{i-1} - 360$$

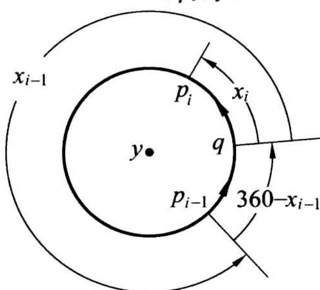


图 22.3

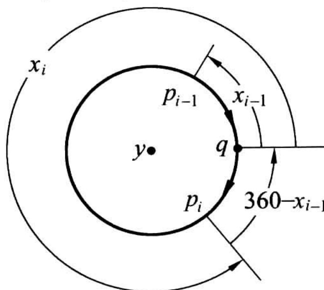


图 22.4

以上三种情况,包括了所有的可能性. 如果相加所有的 $A(\varphi_i, y)$, $i = 1, 2, \cdots, m$, 每一项将有一个 $x_i - x_{i-1}$,其中 r 项将各有一个 $+360$,而 s 项将各有一个 -360. 所以

$$A(\varphi, y) = (x_m - x_{m-1}) + (x_{m-1} - x_{m-2}) + \cdots + (x_1 - x_0) + r360 - s360 = x_m - x_0 + (r - s)360$$

这就完成了证明.

系 如果 φ 是一闭曲线,则 $A(\varphi, y) = (r - s)360$.

只要注意,对于闭曲线,$\varphi a = \varphi b$,所以 $x_m = x_0$,就可得到系的证明.

现在我们已达到最后的阶段来给出围绕数的一个精确定义

$$W(\varphi, y) = \frac{A(\varphi, y)}{360} = r - s$$

从而围绕数是一个整数.

23 $A(\varphi,y)$ 与 $W(\varphi,y)$ 的性质

在精确地定义了 $A(\varphi,y)$ 与 $W(\varphi,y)$ 之后，现在应该证明它们确有我们声称的性质.

1. 如果 φ 是一个点曲线，则

$$A(\varphi,y) = 0 \text{ 与 } W(\varphi,y) = 0$$

因为 $\varphi[a,b]$ 是一个点，φ 是短曲线，平凡的划分是充分地细的. 但对于短曲线，$\varphi a = \varphi b$ 蕴含着 $x_a = x_b$，故 $A(\varphi,y) = 0$；因而 $W(\varphi,y) = \dfrac{A(\varphi,y)}{360} = 0$.

2. $A(\varphi,y)$ 具有关于 φ 的可加性. 明确地说，设 $a < b < c$，并且 $\varphi: [a,c] \to P$. 令 $\varphi_1 = \varphi \mid [a,b]$，又 $\varphi_2 = \varphi \mid [b,c]$，则 $A(\varphi,y) = A(\varphi_1,y) + A(\varphi_2,y)$. 如果 $\varphi a = \varphi b = \varphi c$，$\varphi_1,\varphi_2,\varphi$ 都是闭曲线，而

$$W(\varphi,y) = W(\varphi_1,y) + W(\varphi_2,y)$$

令 $\wp_1$ 及 $\wp_2$ 分别为 φ_1,φ_2 的充分细划分，则 $\wp_1,\wp_2$ 诸顶点的并集给出 φ 的一充分细的划分. 因为 $A(\wp,y)$ 的各项都与 $A(\wp_1,y) + A(\wp_2,y)$ 的各项相同，显然

$$A(\wp,y) = A(\wp_1,y) + A(\wp_2,y)$$

这就证明了第一个关系. 在 $\varphi,\varphi_1,\varphi_2$ 为闭曲线时，这关系的各项是 360 的整数倍. 除以 360，就给出围绕数的关于 φ 的可加性.

24 曲线的同伦

在下一节中将要证明：曲线在一点处的围绕数不会改变，如果这曲线或这点依某种方式连续地变化（见第 18 节）. 本节的目的是明确地叙述我们将容许的变化种类.

定义 令 φ_0 与 φ_1 为空间 Y 中两条曲线，它们的定义域是同一个区间 $[a,b]$，则 φ_0 到 φ_1 的同伦是一矩形 Q 到 Y 的映射，使 Q 的底边映成曲线 φ_0，而顶边映成 φ_1. 明确地说，令 Q 为两变量 (t,τ) 平面中的矩形：$a \leqslant t \leqslant b$ 及 $0 \leqslant \tau \leqslant 1$，则 φ_0 到 φ_1 的一个同伦 Φ 是一个映射 $\Phi: Q \to Y$，使得

$$\Phi(t,0) = \varphi_0 t \text{ 与 } \Phi(t,1) = \varphi_1 t, \text{对于所有的 } t \in [a,b]$$

在 φ_0,φ_1 都是闭曲线时，作为闭曲线 φ_0 到闭曲线 φ_1 的同伦是上述的同伦 Φ，再满足下述条件

$$\Phi(a,\tau) = \Phi(b,\tau) \quad \text{对于所有 } \tau \in [0,1]$$

这个定义有极透彻的几何意义. 画出由水平线段 s_τ 组成的矩形 Q，每一值 $\tau \in [0,1]$ 有一条 s_τ. Φ 在线段之一 s_τ 上的限制决定一由 $\varphi_\tau t = \Phi(t,\tau)$ 所定义



的曲线 $\varphi_\tau:[a,b]\to Y$. 于是得到一族曲线；对应于0与1之间每一个值 τ，有一条曲线(图24.1). 如果把 τ 看做时间变量，那么就可把这族曲线看成是单独一条动曲线的各种位置. Q 中每个铅垂线段映成动曲线的一点所描出的道路. 因为这个图形，一个同伦经常也叫做一个形变.

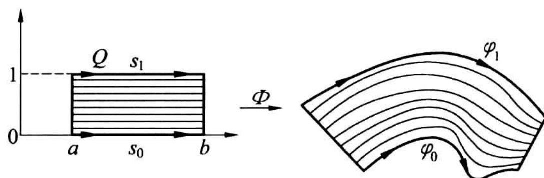


图24.1

如果两曲线 φ_0,φ_1 映射$[a,b]$到平面 P 或到 $\mathbf{R}^m$，则有从 φ_0 到 φ_1 的一种特别的同伦，叫做线性同化. 对于 Q 中各对(t,τ)，定义线性同伦 $\Phi(t,\tau)$ 为把从 $\varphi_0 t$ 到$\varphi_1 t$ 的线段以比$\tau:1-\tau$ 分割的分点(图24.2). 比0 : 1给出线段的起点，而1 : 0给出终点；故 $\Phi(t,0)=\varphi_0 t$ 而$\Phi(t,1)=\varphi_1 t$. Φ 在Q 的各竖直线段上的限制是一相似映射，因为比的保持是相似映射的特征①.

图24.3给出线性同伦的说明. 图中已画出动曲线在时刻 $\tau=0,\frac{1}{4},\frac{1}{2},\frac{3}{4}$ 与1时的位置；以及动曲线上对应于 $t=0,\frac{1}{6},\frac{1}{3},\frac{1}{2},\frac{2}{3},\frac{5}{6}$ 与1的各点所描出的直线道路. 注意，动曲线上单独一点总以常速度运动，走出一条直线段.

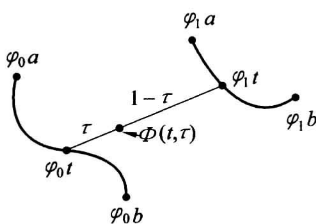


图24.2

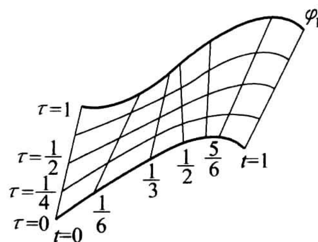


图24.3

必须证明线性同伦 Φ 连续②. 用(c,γ)表示 Q 的一点的坐标. 我们要证明 Φ 在该点处连续，令(t,τ)表示 Q 的任一其他点的坐标. 引进缩写

① 所以，用一点解析几何知识，可知有在 P 或 $\mathbf{R}^m$ 中的式子

$$\Phi(t,\tau)-\varphi_0 t:\varphi_1 t-\Phi(t,\tau)=\tau:1-\tau$$

从而又推导出

$$\Phi(t,\tau)=(1-\tau)\varphi_0 t+\tau\varphi_1 t$$

这里因为 t 已给定，从而 $\varphi_0 t$ 与$\varphi_1 t$ 是已知的两点. 所以这后一个式子是连续 $\varphi_0 t$ 与 $\varphi_1 t$ 这两点的直线段的以 τ 为参数的参数表示.——译者注

② 对于熟悉向量代数的读者，我们能写 $\Phi(t,\tau)=(1-\tau)(\varphi_0 t)+\tau(\varphi_1 t)$，并能论证$(1-\tau)(\varphi_0 t)$是连续的，因为它是连续数量函数$1-\tau$与连续向量函数$\varphi_0 t$的乘积. 同样，$\tau(\varphi_1 t)$连续. 最后，$\Phi(t,\tau)$是连续的，因为它是两个连续函数的向量和.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

$$u = \varphi_0 c, v = \varphi_1 c, z = \Phi(c,\gamma), z'' = \Phi(c,\tau)$$
$$u' = \varphi_0 t, v' = \varphi_1 t, z' = \Phi(t,\tau)$$

图 24.4 显示(t,τ)接近(c,γ)时的情形.实线段 uu' 与 vv' 为:当 τ 从 0 变到 1 时,点 φc 与 φt 在同伦下描出的道路.我们要证明:把(t,τ)限制于(c,γ)附近,就能使距离 $d(z,z')$ 变小(小于一指定的 $\varepsilon > 0$).由三角不等式

$$d(z,z') \leqslant d(z,z'') + d(z'',z')$$

因为从$[0,1]$到线段 uv 的映射(它把 τ 映成$\Phi(c,\tau)$)是相似映射,它是连续的.因此,对应于正数$\frac{\varepsilon}{2}$,有一 $\delta' > 0$,使得对于满足 $|\tau - \gamma| < \delta'$ 的每个τ

$$d(z,z'') < \frac{\varepsilon}{2}$$

因为 φ_0,φ_1 都在 c 处连续,存在着数 $\delta_0 > 0$ 与 $\delta_1 > 0$ 使得

$$d(u,u') < \frac{\varepsilon}{2},\text{对于满足 } |t - c| < \delta_0 \text{ 的每个 } t$$

$$d(u,v') < \frac{\varepsilon}{2},\text{对于满足 } |t - c| < \delta_1 \text{ 的每个 } t$$

现令 δ 为δ',δ_0与 δ_1 中最小者.如果(t,τ)在 $N((c,\gamma),\delta)$ 中,则所有前述三个不等式都成立.还能证明 $d(z'',z')$ 不大于 $d(u,u')$ 与 $d(v,v')$ 二者中的较大者.图 24.4 显示出,当 $d(u,u')$ 是较大者时,如何证明这个事实.通过 v' 作一平行于 vu 而与 uu' 相交于s 的直线,通过 z' 作一平行于uu' 而交$v's$ 于w 的直线.于是直线 qw 延长后交vu 于一点r.由相似三角形,r 分vu 的比,等于 z' 分$v'u'$ 的比;因此 $r = z''$.于是

$$d(z'',z') \leqslant d(z'',w) + d(w,z') \leqslant$$
$$d(u,s) + d(s,u') = d(u,u')$$

所以 $d(z'',z') < \frac{\varepsilon}{2}$.把这与 $d(z,z'') < \frac{\varepsilon}{2}$ 以及三角不等式组合起来,得 $d(z,z') < \varepsilon$.这就完成了 Φ 连续的证明.

在两曲线 φ_0,φ_1 都是闭曲线时,φ_0 到 φ_1 的线性同伦仍然是:φ_0 与 φ_1 为闭曲线时的同伦;因为 $\varphi_0 a = \varphi_0 b$ 与$\varphi_1 a = \varphi_1 b$ 蕴涵从$\varphi_0 a$ 到$\varphi_1 a$ 以及从$\varphi_0 b$ 到$\varphi_1 b$ 的线段重合,从而对于所有的 $\tau \in [0,1]$,$\Phi(a,\tau) = \Phi(b,\tau)$.

当同伦的终端曲线 φ_1 是一条常数曲线即单独的一点时,我们就说该同伦把初始曲线 φ_0 收缩成一点.这种同伦的一个重要例子如下.设 D 为以z 为中心,以圆 C 为边界的圆片.设 $\varphi_0:[0,1] \to C$ 是C 作为闭曲线的标准表示;即$[0,1]$恰好按逆时针方向包围 C 一次(见 16 节).设 $\varphi_1:[0,1] \to z$ 为在中心的常函数所给出的曲线.最后设 Φ 是从φ_0 到 φ_1 的线性同伦,则 Φ 收缩φ_0(或 C)成一点.如把同伦画成一动曲线,则在各瞬时它都是以 z 为中心的圆,并且各点沿径向运动趋于 z.



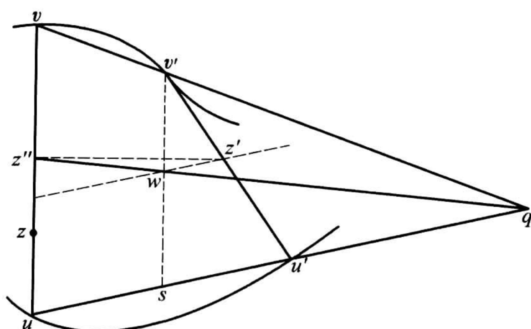


图 24.4

现在设 $f:D\to P$ 是从圆片到平面的一映射，又设 Φ 是方才所描述的同伦. 则复合 $f\Phi:Q\to P$ 是一收缩，它把闭曲线 $f\varphi_0:[a,b]\to P$ 收缩成常的闭曲线 $f\varphi_1$，即点 fz. 这个同伦给出主要定理直观证明中所用的闭曲线族(见第 18 节).

25 围绕数的常值性

定理 25.1 设 $\Phi:Q\to P$ 为从闭曲线 φ_0 到闭曲线 φ_1 的同伦，又设 y 为不在象 ΦQ 中的一点. 则当 τ 从 0 变到 1 时，围绕数 $W(\varphi_\tau,y)$ 是常数. 特别地，有 $W(\varphi_0,y)=W(\varphi_1,y)$.

把 $W(\varphi_\tau,y)$ 缩写为 $f\tau$；f 是定义在 $[0,1]$ 上的函数，根据第 22 节中的证明，f 的每一个值是一个整数. 证明的主要部分是指出：在 τ 的"微小变化"下，f 是常数. 明确地，如果 $\alpha\in[0,1]$，则存在 α 的一邻域 N_α，使得对于每个 $\tau\in N_\alpha$，有 $f\tau=f\alpha$. 一旦证完这事实，就可证明定理如下：因为在 N_α 中 f 是常数，它在 N_α 中连续，从而在 α 处连续；又因为这对于每一个 α 值都正确，故 f 在 $[0,1]$ 中连续. 现在用反证法. 设 f 不是常数并且至少有两个不同的值，则在一维中的主要定理断言：f 取这两个值之间的所有值，包含其间的非整数值. 这与 f 的每一个值是整数的事实相矛盾，故 f 必为常数.

为了证明邻近于 $\alpha\in[0,1]$ 时的常值性，对于 φ_α，选取一充分细的划分 $\wp$. 首先要找出 α 的一邻域 N' 使得，对于所有的 $\tau\in N'$，$\wp$ 对于 φ_τ 是充分地细的. 然后，利用第 22 节中所述的计算围绕数的方法，要找一个更小的邻域 N，使得在 N 中计算的每一个个别步骤保持常值性.

令 $t_0,t_1,\cdots,t_m$ 为 $\wp$ 的顶点. 对每个子区间

$$I_k=[t_{k-1},t_k]$$

$\wp$ 的细，指的是：有一从 y 开始而不交 $\varphi_\alpha I_k$ 的射线 L_k. 令 D_k 表示 Q 中从 (t_{k-1},α) 到 (t_k,α) 的线段. 我们将证明有一包含 D_k 的矩形 E_k，使得它的象 ΦE_k 不与射线

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

L_k 相交.令

$$V_k = \Phi^{-1}(P - L_k)$$

因为 Φ 连续,并且 $P - L_k$ 是开集,从而 V_k 是开集.因为 $\Phi D_k = \varphi_\alpha I_k$ 是在 $P - L_k$ 中,从而 $D_k \subset V_k$.对于每一点 $p \in D_k$,可选一圆形邻域 $N(p) \subset V_k$.尔后,在 $N(p)$ 中取其边平行于(t,τ)轴的最大方形,令 $M(p)$ 表示这方形的内部.对于所有的 $p \in D_k$,这些 $M(p)$ 的集合是 D_k 的一开覆盖.因为 D_k 是紧致的,这覆盖包含一有限覆盖,设为 $M_1, M_2, \cdots, M_n$.令 δ_k 表示这些方形中最小一个的宽度的一半,又令 E_k 表示(t,τ)值的矩形,使得 $t \in I_k$ 与

$$|\tau - \alpha| < \delta_k$$

由于 δ_k 的这样选择,有

$$E_k \subset M_1 \cup M_2 \cup \cdots \cup M_n \subset V_k$$

这是说:对每个 $\tau \in N(\alpha, \delta_k)$,$\varphi_\tau I_k \subset P - L_k$.假定对于每一区间 $I_k, k = 1, 2, \cdots, m$,已作出 $N(\alpha, \delta_k)$;令 δ 为这些数 $\delta_1, \delta_2, \cdots, \delta_m$ 中最小的一个,又令 $N' = N(\alpha, \delta)$.那么,如果 $\tau \in N'$,对于所有的 $k = 1, 2, \cdots, m$,就有 $\varphi_\tau I_k \subset P - L_k$.这就证明,对于每个 $\tau \in N'$,$\wp$ 对于 φ_τ 是充分地细的.

我们现在把第 22 节中计算围绕数的方法应用于这划分 $\wp$ 以及各种曲线 $\varphi_\tau, \tau \in N'$.如同第 22 节,令 C 表示以 y 为中心而半径为 1 的圆.令 $g: P - y \to C$ 表示从 y 到 C 的径向射影,则 $g\Phi: Q \to C$ 与 $g\varphi_\tau: [a, b] \to C$ 都连续,因为它们都是连续函数的复合.在 $g\varphi_\tau$ 作用下,$\wp$ 的顶点 t_k 在 C 上的象,即 $g\varphi_\tau t_k$,缩写为 $p_k\tau$.令 $A_k\tau$ 表示 C 上从 $p_{k-1}\tau$ 到 $p_k\tau$ 的弧,它不与 L_k 相交.如果 q 是 C 上不同于 $p_0\tau, p_1\tau, \cdots, p_m\tau$ 的一点,则如同第 22 节,它的围绕数 $W(\varphi_\tau, y)$ 是 $f\tau = r - s$,其中 r(相应地 s)是包含 q 的正(负)指向的弧的个数.

挑选一点 $q \in C$,不同于 $p_0\alpha, p_1\alpha, \cdots, p_m\alpha$.对于各 $k = 1, 2, \cdots, m$,选取 C 上 $p_k\alpha$ 的一邻域 U_k,它不包含 q 并且与 L_k 或 L_{k+1} 都不相交,则 U_k 将是 C 的包含 $p_k\alpha$ 的一个短弧.因为 $g\Phi$ 连续而 U_k 是 $p_k\alpha$ 的一邻域,并且 $g\Phi(t_k, \alpha) = p_k\alpha$,从而有 α 在$[0,1]$中的一邻域 N_k,使得对于所有的 $\tau \in N_k$,$g\Phi(t_k, \tau) = p_k\tau$ 在 U_k 中.令 N 表示诸邻域 N' 以及 $N_1, N_2, \cdots, N_m$ 中最小的一个,则对于每一个 $k = 1, 2, \cdots, m$,$\tau \in N$ 蕴含 $p_k\tau \in U_k$.(注意:$p_m\tau = p_0\tau$,因为每一个 φ_τ 是闭曲线)特别是,q 与 $p_0\tau, p_1\tau, \cdots, p_{m-1}\tau$ 都不同,因为 q 不在这些 U_k 的任一个之中.

剩下来还要证明:对于各 $\tau \in N$ 以及各 $k = 1, 2, \cdots, m$,两个弧 $A_k\tau$ 与 $A_k\alpha$ 对于 q 具有相同关系;于是对于 φ_τ 同 φ_α,得同一个计数 $r - s$.首先考虑 k,它使 U_{k-1} 与 U_k 有一共同点的情况.则 $U_{k-1} \cup U_k$ 为一连通弧,并且它在 C 中的余弧 D 为一含 q 的与 L_k 相交的连通弧,但对于所有的 $\tau \in N$,即不包含 $p_{k-1}\tau$ 又不包含 $p_k\tau$.因为 D 是连通的,它就全部处于 C 上从 $p_{k-1}\tau$ 到 $p_k\tau$ 的两弧中的一个之





中;其中的一个是 $A_k\tau$,而另一个与 L_k 相交.因为 D 与 L_k 相交,从而 D 不与 $A_k\tau$ 相交.因为 q 在 D 中,这就证明了,对所有的 $\tau \in N$, $A_k\tau$ 不包含 q.于是,在这种情况,$A_k\tau$ 对 q 的关系,对于 $\tau \in N$,是常数;也就是,若 $A_k\alpha$ 不含有 q,则对于 $\tau \in N(\alpha)$, $A_k\tau$ 就不含有 q.

最后,考虑一个 k,它使 U_{k-1} 与 U_k 无公共点.则在 C 中,$U_{k-1} \cup U_k$ 的余集包含两个弧 D 与 E,而令 D 表示与 L_k 相交的那个弧.如上面所论证的,可见对于所有的 $\tau \in N$, D 不与 $A_k\tau$ 相交.此外,因为 E 不与 L_k 相交,对于 $\tau \in N$,所有弧 $A_k\tau$ 包含 E,并且同样地从 U_{k-1} 指向 U_k.现在 q 不在 $U_{k-1} \cup U_k$ 中,所以它一定在 $D \cup E$ 中.如果 $q \in D$,就没有弧 $A_k\tau$, $\tau \in N$,包含 q.如果 $q \in E$,所有弧 $A_k\tau$, $\tau \in N$,包含 q 并且是同样的转向.于是当 τ 在 N 中变化时,$A_k\tau$ 对 q 的关系是常数.这就对于所有的 $\tau \in N$,证明了围绕数 $f\tau$ 的常值性.再用紧接在定理25.1后的论点,就证明了在曲线的同伦下围绕数的常值性.

定理 25.2 设 $\varphi:[a,b] \to P$ 为平面中一曲线.设 y_0, y_1 为 P 的不在曲线 φ 上的两点,并且存在一曲线 $\psi:[0,1] \to P$,它连接 y_0 与 y_1,但不与 φ 相交,则有围绕数的等式

$$W(\varphi, y_0) = W(\varphi, y_1)$$

对于各个 $t \in [a,b]$ 与 $\tau \in [0,1]$,作向量以 φt 为起点,并且与从 $\psi\tau$ 到 y_0 的向量平行,长度相等.令 $\Phi(t,\tau)$ 为这新作的向量的终点(图 25.1).于是,如果对于一固定的 τ,令

$$\varphi_\tau t = \Phi(t,\tau)$$

则曲线 φ_τ 可由曲线 φ 的平移而得到.想象平面 P 为一刚硬的金属片,带有沿着曲线 ψ 切割出来的一条槽沟.用一直径小于槽沟宽度的钉子,在金属片 y_0 处,把该片钉在墙上.现在把金属片贴着墙滑动使得钉子沿着槽沟走,但不容许金属片转动.闭曲线的运动结果,给出了所建立的同伦 Φ 的图形.

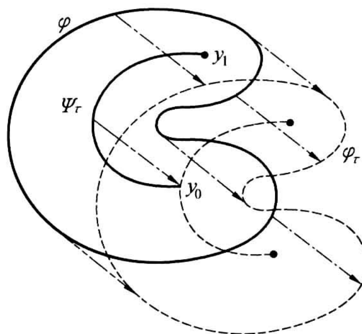


图 25.1

根据在同伦下围绕数的常值性,有

$$W(\varphi_0, y_0) = W(\varphi_1, y_0)$$

此外,在从 y_0 到 y_1 这向量所给出的平面平移下,(φ_1, y_0) 这一对完全叠合(Congruent)于 (φ_0, y_1) 这一对.因为围绕数显然不会因叠合而改变,故 $W(\varphi_1, y_0) = W(\varphi_0, y_1)$.合并这些等式就给出这定理的结论.

26 主要定理的证明

在第 18 节中所叙述的主要定理,现在我们手边有了证明该定理所需要的一切工具.设 y 不在象 fD 中.令 $\varphi_0:[0,1]\to C$ 作为一闭曲线 C 的标准表示法.令 Φ 为第 24 节所述的同伦,它将 φ_0 在 D 上面收缩到 D 的中心点 z.则 $f\Phi$ 是 $f\varphi_0$ 成为在 fz 处常曲线的同伦.因 $f\Phi Q\subset fD$,y 不在同伦的象中.所以由定理 25.1,$W(f\varphi_0,y)=W(f\varphi_1,y)$.因 $f\varphi_1$ 是常曲线,$W(f\varphi_1,y)=0$(见第 23 节),因此 $W(f\varphi_0,y)=0$.于是证明了:如 y 不在 fD 中,则 $W(f\varphi_0,y)=0$.所以 $W(f\varphi_0,y)\neq 0$ 蕴含着 y 在 fD 中.这就完成了证明.(图 26.1 说明对于第 15 节所述的映射 f,同伦 $f\Phi$ 的相继阶段)

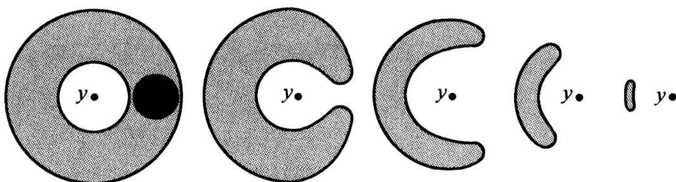


图 26.1

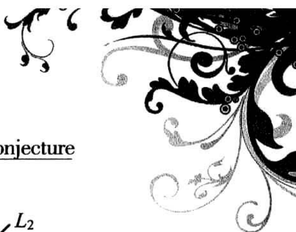
27 圆在各内点处的围绕数是一

在本节中我们将证明:如果一个从圆片 D 到一平面的映射保持圆片边界的每一个点都不动,则 D 的所有点都在 D 的象中.为了铺平证明这一定理的道路,我们首先来证明在第 17 节中根据直观认为明显的下述命题:圆在各内点处的围绕数是一.

引理 若 C 是平面中的一个圆,y 是 C 内部的任一点,并且 $\varphi_0:[0,1]\to C$ 是把 C 看成为一闭曲线的标准表示法(看第 16 节),则

$$W(\varphi_0,y)=1$$

为了证明这引理,采用划分 $\mathscr{P}=\left\{0,\frac{1}{2},1\right\}$.回忆 φ 是这样定义的:选取一参考点 $\varphi 0$,于是 φ 把各个 $t\in[0,1]$ 映成 C 的点 φt,从 $\varphi 0$ 到 φt 的弧所对的圆心角是 $360t°$.所以 $\varphi\left[0,\frac{1}{2}\right]$ 是逆时针方向从 $\varphi 0$ 到 $\varphi\frac{1}{2}$ 的半圆,$\varphi\left[\frac{1}{2},1\right]$ 是逆时针方向从 $\varphi\frac{1}{2}$ 到 $\varphi 1=\varphi 0$ 的半圆(图 27.1).取 y 为中心,从第 22 节的规定得 $W(\varphi,y)=r-s$,其中 $r=1,s=0$.于是在 y 是中心时,引理成立.但各内点都能用不与 C 相遇的线段与中心连接起来.因而,由定理 25.2,它与中心有相同



的围绕数.

考虑把一薄而可伸缩的橡皮圆片的圆边粘在一桌面上的效果,作为下一命题的直觉的提法.保持橡皮圆片的边不动,同时伸长、拉扯或扭弯这橡皮圆片;如果想要看见圆片下面是什么,我们会什么也看不见.

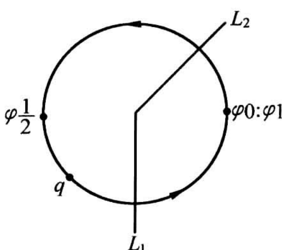


图 27.1

定理 27.1 设 $f:D\rightarrow P$ 是从一圆片 D 到平面 P 的一个映射,它保持边界圆 C 的各点不动,则象 fD 包含 D 的全部.

令 $\varphi_0:[0,1]\rightarrow C$ 为 C 的作为一闭曲线的标准表示,又令 y 为 D 的一内点.因 f 保持 C 的各点不动,$f\mid C$ 是恒同映射,就是 $f\varphi_0=\varphi_0$.所以 $W(f\varphi_0,y)=W(\varphi_0,y)$.前面的引理断言 $W(\varphi_0,y)\neq 0$;因而 $W(f\varphi_0,y)\neq 0$.主要定理现在适用并断言 $y\in fD$.C 的各点在 D 中,因 $fC=C$.于是 D 的全部在 fD 中.

作为本定理的一个推论,有

定理 27.2 不存在从圆片 D 到它的边界 C 的,而保持 C 的各点不动的连续映射.

人们当然能把一矩形及其内部映射成它的边界中的一条边,使得这条边保持固定,设想把挂在墙上的一幅国画取下卷起来;在这映射 f 的作用下,对于画轴上的全部点 x,$fx=x$.由画轴所表示的边,称为矩形区域(国画)的一个收缩核.推论说的是我们不能把一个圆的内部卷起来成为它的边界;就是说,一个圆不是圆片的一个收缩核.想象圆及其内部好比是鼓的边缘与鼓膜.断言是:整张鼓膜无法拉开而卷成边缘.人们根据这形象化的例子,有时把这个推论称作鼓面原则.

28 不动点性质

在第 9 节中,我们证明了一条线段的任一自映射至少有一个不动点.现在我们要对于一个圆片证明类似定理.

定理 28.1 令 $f:D\rightarrow D$ 为一个圆片 D 的自映射,则 f 至少保持 D 的一个点不动,即存在至少一个点 $x\in D$ 使得

$$fx=x$$

假设不然,即假设存在一个无不动点的映射 $f:D\rightarrow D$,则对于各个 $x\in D$,fx 与 x 都不同.因而可以作一条射线 L_x,从 fx 出发而且通过 x.令 gx 为 L_x 与 C 的交点.如果 fx 在 C 上,gx 是 L_x 与 C 相交的另一点;又如果 $x\in C$,则 $gx=x$.图 28.1 表明可能性中的几种.于是有 $g:D\rightarrow C$,与 $g\mid C$ 是恒同变换.我们要证

明 g 连续；然后 g 与定理 27.2 相矛盾，而这矛盾将证明我们的定理.

图 28.1

要证明 g 的连续性，令 $x_0 \in D$，又令 V 为 gx_0 的一邻域. 我们来作 x_0 的一邻域 U 使得 $x \in U$ 蕴含 $gx \in V$. 令 b, c 表示 V 的端点，又令 m 为线段 x_0 到 fx_0 的中点. 令 H 为通过 b 与 m 的直线，又令 K 为通过 c 与 m 的直线. 选取 fx_0 的一个圆形邻域 N，使它不包含 H 或 K 的点. 因 f 连续，就有 x_0 的一个圆形邻域 U' 使 $fU' \subset N$. 现在选取 x_0 的一邻域 U，使 U 不包含 H 或 K 的点，并且 $U \subset U'$. 则我们有如图 28.2 所示的情形：U 是 x_0 的一邻域，N 是 fx_0 的一邻域，U 与 N 都不与 H 或 K 相遇，以及 $fU \subset N$. 此外，一点 $x \in U$ 以及它的象 $fx \in N$ 位于 H（并且也是 K）的两侧，因 x_0 与 fx_0 位于 H（并且也是 K）的两侧，并且 U 与 N 都是不与 H（以及 K）相遇的连通集.

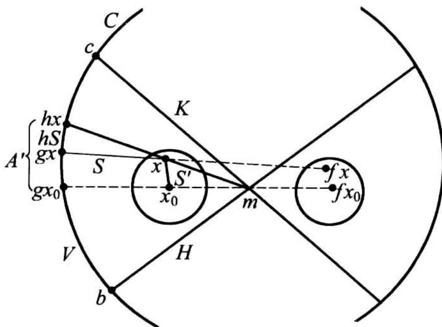


图 28.2

因此从 fx 出发经过 x 的射线 L_x，与 H 以及 K 两者都在 fx 与 x 间相遇. 所以 L_x 上从 x 到 gx 这线段不包含 H 或 K 的点. 令 h 表示从 m 映到 C 的径向射影. 因 m 在 L_{x_0} 上，h 把 x_0 映成 gx_0. 现在 h 把从 x_0 到 x 这线段 S' 映成 C 的从 gx_0 开始的一条弧 A'. 因为 S' 不与 H 或 K 相遇，A' 不能包含 b 或 c. 因而 A' 完全在 V 中，故 $hx \in V$. 因为 S 不与 H 或 K 相遇，与前面一样，随即有 $hS \subset V$. gx 的从 m 到 C 的径向射影是 gx，即 $hgx = gx$. $gx \in S$，$gx = hgx$，$hS \subset V$ 这三个断言的结合蕴涵 $gx \in V$，而这就完成了 g 的连续性的证明.



29 向量场

平面或空间中的一个向量是一对有顺序的点.习惯上用连接两点的一条有向线段来表示一个向量,其方向用从第一点指向第二点的箭头表明.向量的代数性质使得向量成为研究高维欧氏空间几何必不可少的工具.在数学物理中,用向量来表示力、速度以及加速度,也非常重要.

我们将要运用速度向量这一概念,来帮助我们直观地领会下面几节中所要证明的定理.如果一动点通过一点 x,它在 x 处的速度向量是以 x 为起点、以运动的瞬时方向为指向而长度为瞬时速率的这个向量 $\boldsymbol{v}$.如果这点以同一速率在同一方向继续运行一单位时间,它就会到达 $\boldsymbol{v}$ 的终点 x'.常速(即常方向与常速率)只是最简情况,但我们必须考虑动点沿一曲线的运动.当时,当点沿曲线移动时,它的方向与速率通常会改变.曲线每一点处的速度向量切于这曲线,指向与曲线的方向相同,它的长度是瞬时速率((弧长)/(时间)在时间趋近于0时的极限).例如,设一质点沿一曲线运动,在 t 单位时间时,它在曲线的点 t 处,$t=0,1,2,3,\cdots$(图29.1左);而附在各点的切向量表示质点在这些点处的速度(图29.1右).于是,在点1与2处较小的速度向量,符合于从1到2以及从2到3较短的距离;其他各点处的切向量表明点3处的速度突增,点4处减速而在点5处高速,等.

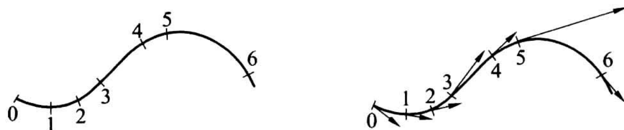


图 29.1

向量场 $\boldsymbol{v}$ 是一个函数,它对于平面(或空间)一个区域的每一点 x,指定一个以 x 为起点的向量 $\boldsymbol{vx}$.如果考虑流动的液体或气体,不同质点在单独一瞬时的速度向量形成一个向量场.例如在一个以不变速度和定常方向的稳定流体中,向量都平行而等长.作为另一例,考虑平面绕着一点 z 作常角速度的旋转(图29.2).

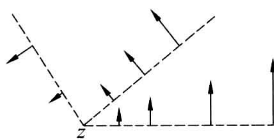


图 29.2

在一点 x 处,$\boldsymbol{vx}$ 垂直于连接 z 与 x 的直线,而其长度与 $d(z,x)$ 成比例.向量 $\boldsymbol{vz}$ 是有序的一对点(z,z);它无方向而长度为零.这样的一个向量叫做零向量.

一流动液体的向量场在所有时刻都相同时,叫做稳定流.明确地说,速度向量仅依赖于质点在平面(或空间)中的位置,而不依赖于时间.在不同时刻通过同一点的两质点在该点有同一速度向量.上面讨论的两例,都是稳定流的例.这

样的流有流线，都是质点的道路．它们形成一曲线族，使得通过每一点仅有一曲线，并且在该点处的速度向量切于曲线．沿任何流线的质点始终在该流线上．能把流线当成是沿它自身的滑动．在上面的第一例中，流线形成一族平行线．在第二例中，它们形成以 z 为中心的一族圆．通过 z 的流线是在 z 处的常曲线．

30　向量场与映射二者的等价

乍一看，要用数学方法来处理向量场，似乎是相当困难的．但是按照下列方式，向量场的概念完全等价于映射的概念．设 $f:A\to P$ 为从平面的一子集 A 到平面的一个映射．令 o 为 P 的一固定参考点，称它为原点．对于每个 $x\in A$，令 $\boldsymbol{vx}$ 为从 x 出发的向量，它与从 o 到 fx 的向量平行且等长．于是对于每个映射 f 规定了一个向量场 $\boldsymbol{v}$．反之，如果在 A 上已给定了一个向量场 $\boldsymbol{v}$，就能定义 f 如下：fx 是以 o 为起点的，并且与 $\boldsymbol{vx}$ 平行而等长的向量的终点．容易看出，向量场与映射二者之间的这个对应是一对一的．

能用两向量等价这概念来弄清这个对应．两个向量称为等价的向量，如果它们平行，等长且有同样的指向．如果其中一个为零向量，则只有两个都为零时才等价．现在如果 $\boldsymbol{v}$ 是一个向量场，x 处的向量 $\boldsymbol{vx}$ 等价于 o 处的唯一一个向量，而这向量由它的终点 fx 唯一确定．反之，如果 $f:A\to P$ 是一映射，我们就能获得对应的向量场 $\boldsymbol{v}$ 如下：定义 $\boldsymbol{vx}$ 为 x 处的，并且等价于从 o 到 fx 这向量的向量．

图 30.1 的左图，画出了半径为 r 的圆 C 上的向量场，由各点处长为 $\frac{1}{2}r$ 并且指向逆时针的切向量组成．图 30.1 的右图，画出了对应的映射的象．这时候 f 是一个相似变换：把 C 收缩到以 o 为心，大小为 C 的一半的圆，并且把 C 旋转了 $90°$．图 30.2 画出了一个外向的，长为 $\frac{1}{2}r$ 的法线场．这时，f 把圆缩小一半，但不旋转．具有等长的内向法线场会给出另一个 f，它是上面的那个 f 再加上绕 o 旋转 $180°$．

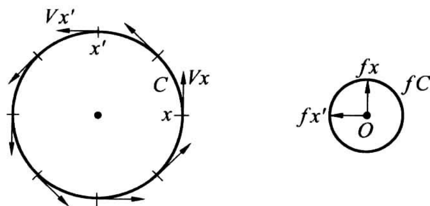


图 30.1

在 f 是常函数，把 A 的全部映到单独的一点时，对于全部 $x\in A$，象 fx 是同一点．于是对于全部 $x\in A$，有序的点对 (o,fx) 都相同，并且对应场的全部向量

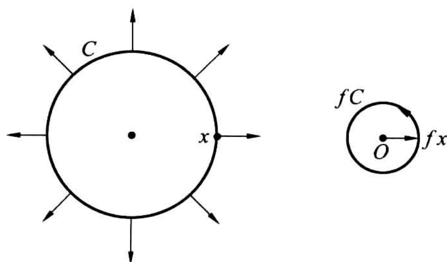


图 30.2

都平行,都有同一长度与同一指向.这样的场叫做常向量场.

因为向量场与映射间的一对一的对应,我们能把为映射而定义的概念和性质转用到向量场.例如,如果一向量场 v 的对应函数 f 连续,我们就称 v 连续.

31 一向量场沿着一闭曲线的指数

令 $\boldsymbol{v}$ 表示平面 P 的子集 A 上的一连续向量场,又令 $\varphi:[a,b]\to A$ 为 A 中的一闭曲线.设想曲线是一动点所描出的.在曲线的各个位置 x 处,向量 $\boldsymbol{vx}$ 有定义.当动点描出这曲线时,向量将连续变化,旋转它的方向并改变它的长度.当这点回到它的初始位置时,向量应恢复到它的初始方向与长度.人们能问:当点走过这曲线时,向量的方向要旋转多少整圈.利用第30节的相应映射 $f:A\to P$,能够非常清楚地陈述这个问题以及它的答案.复合函数 $f\varphi:[a,b]\to P$ 是平面中的一闭曲线.如果原点 o 不在 $f\varphi$ 上,则围绕数 $W(f\varphi,o)$ 有定义.我们把它叫做向量场 $\boldsymbol{v}$ 沿着闭曲线 φ 的指数并且用 $I(\boldsymbol{v},\varphi)$ 表示.于是

$$I(\boldsymbol{v},\varphi)=W(f\varphi,o)$$

在图30.1的例中,显然有 $I(\boldsymbol{v},C)=1$.图30.2的例也一样对,而内向法线的第三例也一样.但是图31.1中的常向量场的指数是零.图31.2给出沿着一圆的指数为2的一向量场.

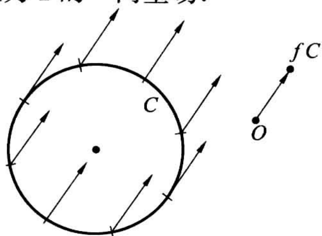


图 31.1

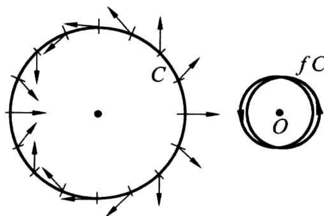


图 31.2

定理31.1 令 $\boldsymbol{v}$ 为定义在平面中一圆片 D 上的一连续向量场,并且使得 D 的边界圆 C 上的任一点 x 处的向量 $\boldsymbol{vx}$ 不是零向量.如果 $\boldsymbol{v}$ 的沿着 C 的指数 $I(\boldsymbol{v}$,

C）不是零，则在 D 中至少有一个点 x，它那里的向量 $\boldsymbol{v}x$ 是零.

把二维时的主要定理的“映射”语言，翻译成“向量场”语言，恰好就是本定理. 令 $f:D\to P$ 为对应于场 $\boldsymbol{v}$ 的映射，并且令 $\varphi:[0,1]\to C$ 为把 C 作为闭曲线的标准表示法. 由假设 $W(f\varphi,o)=I(\boldsymbol{v},\varphi)$ 不是零. 主要定理断言：方程 $fx=o$ 至少有一解 x. 于是对应的 $\boldsymbol{v}x$ 必是零向量，因为它等价于从 0 到 o 的向量.

定理 31.2 令 $\boldsymbol{v}$ 为定义在圆片 D 上一连续的非零向量场. 于是在 D 的边界 C 上，至少有一点 x，在该处的 $\boldsymbol{v}x$ 是一外法线，至少有一点 x'，在该点处的 $\boldsymbol{v}x'$ 是一内法线，并且 C 上至少有两个点，在那些点处的向量都与 C 相切. 一般地，对每个角 α，至少有一点 $x\in C$，使得 $\boldsymbol{v}x$ 与在 x 处的外法线夹角为 α.

这些结论的成立，主要由于 $\boldsymbol{v}$ 在 D 内部非零这一假设. 例如，令 D 的中心为 o，令 $f:D\to P$ 为恒同映射. 对应的 $\boldsymbol{v}$ 正好有一个零在 D 的中心. 定理的结论在这里不成立；因为在 C 的每点 x 处，$\boldsymbol{v}x$ 都是外法线. 常向量场是一个说明这定理的好例子(图 31.1)；各 α 恰好出现一次.

只需要证明定理 31.2 的最后结论，因为它蕴涵前面的各结论. 只要令 $\alpha=0°$ 与 $\alpha=180°$，就分别得到关于外法线与内法线的结论，令 $\alpha=90°$ 与 $270°$ 就得到关于切线的结论.

为证明这最后结论，令 α 固定. 选取 D 的中心为原点 o，又令 $f:D\to P$ 为场 $\boldsymbol{v}$ 的对应映射. 用 g 表示 P 绕着 o 转过角 $-\alpha$ 的旋转. 用 h 表示从 o 到盖满 C 的径向射影. 因 $\boldsymbol{v}$ 永不为零，fD 与 gfD 不会包含 o. 从而 $hgf:D\to C$ 有定义. 因为 $C\subset D$，hgf 能看成为一映射 $D\to D$. 定理 28.1 断言：hgf 至少保持一个点不动；那就是，存在一点 $x\in C$ 使 $hgfx=x$. 令 $\boldsymbol{v}'x$ 为在映射 hgf 作用下对应于 x 的向量. 由定义，它平行于从 o 到 $hgfx=x$ 的向量. 所以 $\boldsymbol{v}'x$ 是在 x 处的外法线. 但对于 D 的任一点 y，$\boldsymbol{v}y$ 与 $\boldsymbol{v}'y$ 怎样区别呢？因为我们得到 hgf，是首先运用 g 到 f，然后运用 h 到 gf，从而得到 $\boldsymbol{v}'y$，它是先把 $\boldsymbol{v}y$ 绕着 y 转过角 $-\alpha$，然后把长度改为 C 的半径长. 于是，在 C 上的固定点 x 处，向量 $\boldsymbol{v}x$ 与 $\boldsymbol{v}'x$ 的夹角一定是 α，其中 $\boldsymbol{v}'x$ 已说明是外法线.

推论 如果 $\boldsymbol{v}$ 为圆片 D 上一连续向量场，又如果在 C 上 $\boldsymbol{v}$ 永不与 C 相切（或永不为 C 的法线），则在 D 中 $\boldsymbol{v}$ 至少有一个零.

前面的结果在稳流的研究中是重要的. 向量场的零只发生在流中保持不动的点处. 设圆片上的向量场是这样的向量场，在其边界 C 上，场为内法线. 流体显然在 C 的各点处流入 D. 直观告诉我们流体必在 D 内某处堆积. 因为场没有与 C 相切之处，上述推论告知我们，流体能汇集在流的至少一个不动点处.

32 球到平面的映射



我们说球 S，指的是空间中与一点 z（中心）的距离为一固定正数 r（半径）

的点的全体.如果 $x \in S$,x 的对径点 x' 是连接 x 与 z 的直线与 S 的另一交点,即 S 的一直径交 S 于相对的两点 x 与 x'.

如果用垂直射影 f 把球 S 映到一平面 P,则有一对对径点 x 与 x' 使得 $fx = fx'$,即过 z 而垂直于 P 的直线与 S 的一对交点.奇怪的是,这结果的一部分,对于 S 映到 P 的任何映射都成立.下面类似于定理10.1的定理是数学家K.Borsuk和S.Ulam在1933年发现的.

定理 32.1 球到平面的每一个映射 $f: S \to P$ 把 S 的某对对径点映成同一点;也就是,对于至少一对对径点 x 与 x',$fx = fx'$.

为了证明定理,随意选取 P 中一点 o 作为原点.定义从 S 到 P 的一个对应 $g: S \to P$ 如下:S 的任一点 x 决定了它的对径点 x'(图32.1),并且 f 决定了 fx 与 fx' 以及从 fx 到 fx' 这向量;gx 是等价于这向量的以 o 为起点的那向量的终点.于是 $g: S \to P$ 有性质:对于每个 $x \in S$,gx' 与 gx 关于 o 对称;因为 x 为 x' 的对径点,只要把从 fx 到 fx' 的箭头反转过来,就得到 gx'.只要能证明 g 把 S 的某点映成 o,定理就证明了;因为只有这样,向量与它的反向量才重合.

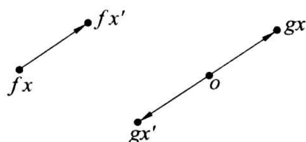


图 32.1

假定已知 g 连续;这在以后会证明的.令 P' 为通过 S 的中心 z 的一个固定平面,令 C 表示 P' 与 S 相交的圆,又令 D 为以 C 为边界的 P' 中的圆片.如果象 gC 包含 o,即存在 S 的一点 x 使得 $gx = o$,定理就证明了.所以现在只需要考虑 gC 不包含 o 的情形.令 H 为 S 被 P' 所分成的两个半球中的一个,又令 $\psi: D \to H$ 为 H 到 D 的垂直射影的逆映射,则复合映射 $g\psi: D \to P$ 在边界 C 上与 g 重合.令 φ 为 C 作为闭曲线的标准表示法(第16节).证明围绕数 $W(g\varphi, o)$ 不是零,本定理就证明了.因为一旦知道围绕数不是零,主要定理(第18节)就向我们保证:D 中存在一点 y 使 $g\psi y = o$,并且这点 y 给出 S 的点 $x = \psi y$ 满足 $gx = 0$.

我们将说明 $W(g\varphi, o)$ 是奇整数(回忆零是偶数,因为 $2 \cdot 0 = 0$),来证明 $W(g\varphi, o) \neq 0$.令 φ_1 与 φ_2 为 φ 在子区间 $\left[0, \frac{1}{2}\right]$ 与 $\left[\frac{1}{2}, 1\right]$ 的限制.令 $x_0 = \varphi 0 = \varphi 1$,则它的对径点 $x'_0 = \varphi \frac{1}{2}$.此外,$\varphi_1$ 表示 C 的一个半圆,作为从 x_0 到 x'_0 的曲线,φ_2 表示另一个从 x'_0 到 x_0 的曲线.现在选取 $\left[0, \frac{1}{2}\right]$ 的一划分,对于曲线 $g\varphi_1$ 关于 o 它是充分细的,并利用第22节的结果来计算 $A(g\varphi_1, o)$,得

$$A(g\varphi_1, o) = u - v + (r - s)360$$

其中 $r - s$ 是整数,而 u 与 v 分别是从 o 到 gx'_0 与从 o 到 gx_0 的两射线在量角器上的读数.因为 x_0, x'_0 是对径点,gx'_0, o, gx_0 这三点都在一直线上;因而 $u - v$

为平角的测量度数，即 $u - v = \pm 180$. 从而 $A(g\varphi_1, o)$ 是 180 的奇数倍

$$A(g\varphi_1, o) = (2m + 1)180$$

图 32.2 表明倍数为 -3 的情形.

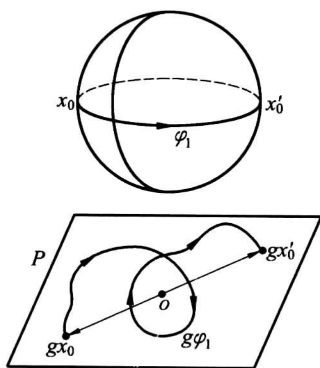


图 32.2

其次考虑从 gx'_0 回到 gx_0 的曲线 $g\varphi_2$. 令 $h: P \to P$ 为绕 o 转过 180° 的旋转. 如果 $t \in \left[0, \frac{1}{2}\right]$，则 $t + \frac{1}{2} \in \left[\frac{1}{2}, 1\right]$，与 $\varphi_2(t + \frac{1}{2})$ 是 $\varphi_1 t$ 的对径点. 从而由 g 的对称性，有

$$g\varphi_2(t + \frac{1}{2}) = hg\varphi_1 t, \text{对所有的 } t \in \left[0, \frac{1}{2}\right]$$

也就是说，曲线 $g\varphi_2$ 是由曲线 $g\varphi_1$ 经过 180° 旋转而得. 因为旋转保持角度，$A(g\varphi_1, o) = A(g\varphi_2, o)$. 运用 A 的可加性质，得

$$\begin{aligned} A(g\varphi, o) = A(g\varphi_1, o) + A(g\varphi_2, o) &= \\ 2A(g\varphi_1, o) &= \\ 2(2m + 1)180 &= \\ (2m + 1)360 \end{aligned}$$

因此
$$W(g\varphi, o) = \frac{A(g\varphi, o)}{360} = 2m + 1$$

这就完成了 $W(g\varphi, o)$ 是奇数的证明.

g 的连续性还待证明. 令 x_0 为 S 的一点，又令 N 为 gx_0 的半径为 r 的圆形邻域. 令 U, U' 分别为 fx_0, fx'_0 的圆形邻域，其半径都为 $\frac{r}{2}$. 因 f 连续，分别有 x_0，x'_0 的邻域 V, V' 使 $fV \subset U$ 与 $fV' \subset U'$. V' 的点的对径点集合 T 是 x_0 的一邻域；令 W 为包含于 V 与 T 两者之中的 x_0 的邻域，于是，若 x 是 W 中一点，我们有 $x \in V$ 与 $x' \in V'$；从而有 $fx \in U$ 与 $fx' \in U'$. 令 y 为 P 的点，使得由 fx_0 到 y 的向量等价于由 fx 到 fx' 的向量(图 32.3). 既然由定义，从 o 到 gx 的向量也等价于从 fx 到 fx' 的向量，而从 o 到 gx_0 的向量等价于从 fx_0 到 fx'_0 的向量，可见距离

$$d(gx, gx_0) = d(y, fx'_0)$$

由三角形不等式，这距离所以小于或等于

$$d(y, fx') + d(fx', fx'_0)$$

从平行四边形，$d(y, fx') = d(fx_0, fx)$，并且这个距离与 $d(fx', fx'_0)$ 两者都小于 $\frac{r}{2}$；所以它们的和小于 r. 这就蕴涵着 $gx \in N$. 因而 g 把 W 映入 N. 这证明了 g 的连续性，于是完成了定理的证明.

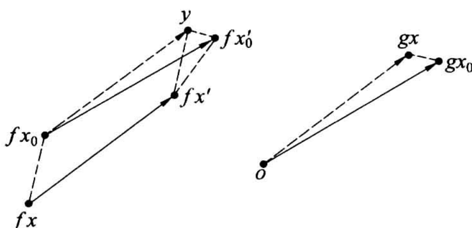


图 32.3

我们来给出一个应用.设地球表面为一球面 S,并且在任何瞬时,气压 px 与气温 tx 都是 $x \in S$ 的连续函数.在平面 P 中,选取一原点,过该点的两条互相垂直的有向直线以及一量度单位,作成直角坐标系.对于 S 的任一点 x,令 fx 为 P 的点,其坐标是 (px, tx).因为 p 及 t 连续,从而 $f: S \to P$ 连续.现在把定理应用到这个 f,就得到

推论 在各瞬时,地球表面有一对对径点,在该两点处的气压相等,并且气温也相等.

显然气压与气温的物理性质都与结论无关;p 与 t 可以是任何两个定义在 S 上的连续实值函数.

也可注意,如果考虑一个单独的函数,例如气温,则定理 10.1 告诉我们,在每一个大圆上有一对对径点,该两点处的气温相等.

33 分火腿三明治

本节定理系定理 11.1 的三维推广.该定理说,平面中任何一对有界连通区域能被单独的一条直线恰好分成两半(面积相等的两半).

定理 33.1 令 A, B, C 为空间中三个有界连通开集,则存在单独的一个平面,它把每一个都恰好分成体积相等的两半.

三个球体以及通过它们中心的平面提供了一个例子,定理的力量在于,即使区域不规则它也成立.如果把 A, B 解释为面包片而 C 为夹在它们之间的火腿,则结论能解释为:能把火腿三明治用一刀切开,使得两片面包与一片火腿都恰好分成两半.

为着证明此定理,先选取包含 A, B, C 的一个球 S.存在着这样一个球,因为 A, B, C 都有界.令球的中心为 z,半径为 r.对于各 $x \in S$,令 L_x 表示过 z 与 x 这直径的直线.我们将证明:

(1) 对于各 $x \in S$,在 L_x 上存在着恰好一个点 x_A,使得在 x_A 处垂直于 L_x 的平面把 A 按体积分成两半.

一旦证明了这一事实,就令 $g_A x$ 为距离 $d(z, x_A)$,它附有正号,如果 x_A 在由

z 到 x 的线段上,而附有负号,如果 x_A 在由 Z 到 x 的对径点 x' 的线段上.因为 L_x 与 $L_{x'}$ 重合,并且在 L_x 与 $L_{x'}$ 所取的坐标系的指向相反,又因为 $x'_A = x_A$,所以

$$g_A x' = - g_A x$$

用 B 与 C 分别替代 A,用类似的方式分别定义 x_B,$g_B x$ 与 x_C,$g_C x$.现在考虑映射 $f: S \to \mathbf{R}^2$,对各 $x \in S$,规定象点的坐标为

$$fx = (g_A x - g_B x, g_A x - g_C x)$$

我们将证明:

(2)f 连续.

一旦证明了(1)与(2),就能很快地完成定理 33.1 的证明如下.由定理 32.1,存在一点 x 使 $fx = fx'$.从 fx 与 fx' 的坐标相等,得到

$$g_A x - g_B x = g_A x' - g_B x'$$

$$g_A x - g_C x = g_A x' - g_C x'$$

再凭借上面已说明的关系 $g_A x' = - g_A x$,$g_B x' = - g_B x$ 以及 $g_C x' = - g_C x$,第一个方程化简为 $g_A x = g_B x$ 而第二个为 $g_A x = g_C x$.因而,对于点 $x \in S$,它的象与它的对径点 x' 的象重合,有 $x_A = x_B = x_C$,并且在这点 x 处垂直于 L_x 的平面把所有三个区域分成两半.

为着证明(1),对各点 $y \in L_x$,令 P_y 表示过 y 而垂直于 L_x 的平面.考虑 A 的与 x 在 P_y 的同一侧的那部分;令 hy 表示它的体积.当 y 从 x' 变到 x 时,hy 从 A 的体积变到零.若 y_1,y_2 都在 L_x 中,差 $|hy_1 - hy_2|$ 至多是实球在两平面 Py_1 与 Py_2 之间的那部分体积,而这是小于 $\pi r^2 |y_1 - y_2|$ 的.这证明 h 在 L_x 的各点 y 处连续(取 $\delta = \frac{\varepsilon}{\pi} r^2$).所以一维时的主要定理保证,存在一点 y 使 hy 是 A 的体积的一半.如果有两个这样的点,那将会有两个平行的平面 P_1 与 P_2 都把 A 分成两半.空间在 P_1 与 P_2 之间的那块厚片 Q,将其余的空间分成两个不连通部分.因为 A 是连通的,并且在各部分有它的一半体积,A 必包含 Q 内部的一点 q.因为 A 与 Q 都是开集,就有 q 的一球形邻域 U 包含在 $A \cap Q$ 之中.既然 U 有正体积,所以 $A \cap Q$ 也如此.从 P_1 移到 P_2,hy 就改变了 $A \cap Q$ 那么大的体积,所以 P_1 与 P_2 都不能分 A 成两半.这就证明了(1).

为着证明(2),f 的连续性,只要证明 f 的各坐标的连续性就够了.我们将只证明 g_A 连续,其余的留给读者.令 c 为 S 的一点,要证明在该点处 g_A 连续,又令 C_A 为 L_C 上的点,在该点处垂直于 L_C 的平面 P_c 切开 A 成两半.令 x 为 S 上挨近 c 的点,又设 x_A,P_A 类似地定义.图 33.1 显示出我们关心的图形与过 c,x,z 的平面的相交.我们要证明,将 x 限制于挨近 c(在 $N(c,\delta)$ 中),能使 $|g_A x - g_A c|$ 小于预定的 $\varepsilon > 0$.

令过 c 与 x 的大圆交 P_c 于 u 与 v 两点.令 P' 与 P'' 分别为过 u 与 v 而垂直



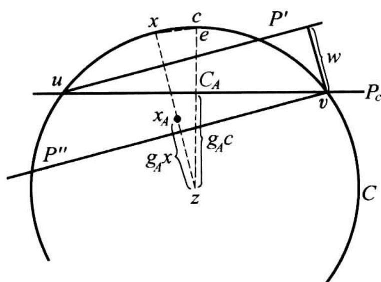


图 33.1

于 L_x 的两平面. 令 N 表示 S 的内部. 与 c 在 P' 同侧的那部分 N 被包含在与 c 在 P_c 同侧的那部分 N 之中. 既然 $A\subset N$, 与 c 在 P' 同侧的 A 的那部分被包含在与 c 在 P_c 同侧的 A 的那部分之中. 所以, 如果以 V 表示 A 的体积, 与 c 在 P' 同侧的 A 的那部分体积至多是 $\frac{V}{2}$. 由类似的论证, 与 c 在 P'' 异侧的 A 的那部分体积, 至多是 $\frac{V}{2}$. 从而 P_x 必在 P' 与 P'' 之间. 所以, 若 w 是 P' 与 P'' 之间的距离, 我们有

$$|g_Ax-g_Ac|<w$$

要得到 w 的大小的估计, 注意从两三角形的相似, 得

$$\frac{w}{d(u,v)}=\frac{d(e,x)}{d(z,x)}$$

其中 e 是 x 在 L_C 上的垂直射影. 因为 $d(z,x)=r$, 即得

$$w=\frac{d(u,v)}{r}d(e,x)$$

因为 $d(u,v)\leqslant 2r$, 并且 $d(e,x)\leqslant d(c,x)$, 得

$$w\leqslant 2d(c,x)$$

所以

$$|g_Ax-g_Ac|<2d(c,x)$$

对于一已知 $\varepsilon>0$, 取 $\delta=\frac{\varepsilon}{2}$, 则 $x\in N(c,\delta)$ 蕴含着 $|g_Ax-g_Ac|<\varepsilon$. 这证明 g_A 在 c 处连续. 因为这对各个 $c\in S$ 成立, 从而 g_A 在 S 上连续. 定理 33.1 的证明到此完成.

34　一球面的切向量场

令 $\boldsymbol{v}$ 表示定义在空间一球面 S 上的一个向量场(见第 29 节). 对于 S 的每一点 x, 它指定了一条以 x 为起点的有向线段. 如果对于 S 的各 x, 以 x 为起点的各线段都切于 S, 或等价地, 垂直于半径 zx, 其中 z 是 S 的中心, 我们就说场 $\boldsymbol{v}$ 是 S 的切向量场. 如同第 30 节中那样, 取一原点 o, 并定义 gx 为从 o 出发而与 $\boldsymbol{v}x$ 平

行且等长的向量的端点，这样来把 S 映入空间的一映射 g 与 v 配合．我们说，当配合的 g 连续时，v 就连续．

定理 34.1　令 v 为在一球面 S 上定义的并且与 S 相切的连续向量场，则 S 上存在至少一个点 x 使得 $\boldsymbol{v}\boldsymbol{x} = \boldsymbol{0}$.

能把 S 的一个切向量场解释为一个流．定理蕴涵：球面上任一稳定流至少有一个静止点．举一个实际例子，设地球表面是一个球面，并且风流的速度向量是连续的．则在任一瞬间，地球上的某一点处无风．

如果把一球绕着一轴以常角速度旋转，我们得到具有两个静止点的流．

为了说明定理，让我们作出 S 的一个切向量场，具有恰好一个零在点 x_0 处．令 L 为 S 的过点 x_0 的一有向切线．在任一不同于 x_0 的点 $x \in S$，由 x 与 L 决定一平面 P_x，它交 S 于圆 C_x．给予 C_x 以相同于其切线 L 的指向．定义 $\boldsymbol{v}\boldsymbol{x}$ 为 P_x 中一向量，它以 x 为起点并且切于 C_x，长为距离 $d(x, x_0)$ 之半，而指向与 C_x 一致．图 34.1 显示 P_x 中若干个切于 C_x 的向量；我们已把 L 的指向向上，于是 C_x 逆时针方向．注意到当 x 趋近于 x_0 时向量越来越短．规定 $\boldsymbol{v}\boldsymbol{x}_0 = \boldsymbol{0}$ 来完成 v 的定义而获得一连续场．

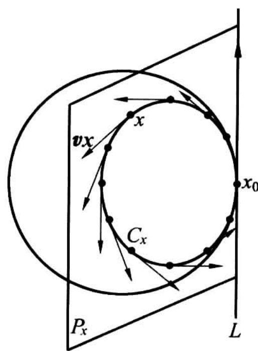


图 34.1

与一球面形成对照，环面（车轮内胎）具有无处为零的、连续的切向量场．画出车轮内胎绕一轴旋转的速度场．同样，画出一烟圈．

证明　选定一个过 S 的中心 z 的平面 P．令 C 表示圆 $P \cap S$，又令 D 为 P 中心 C 为边界的圆片．用 H 与 H' 表示 S 的由 C 所决定的两个闭半球．考虑过 z 并垂直于 P 的直线；令它与 H 及 H' 的交点分别为极点 p 及 p'．令 $h: H' \to D$ 为从 p 作球极射影所给定的拓扑等价．明确地说，如果 $x \in H'$，则 hx 是 D 与从 p 到 x 这线段的交点．类似地，令 $h': H \to D$ 为从 p' 作球极射影所给定的拓扑等价．

只要能证明下述断言就够了：如果 $\boldsymbol{v}$ 在两个半球之一上无零向量，譬如说在 H' 上，则在 H 上必有一个零向量．作为第一个主要步骤，我们将证明：

(1) 球极射影 h 把 H' 上的场 $\boldsymbol{v}$ 映入 D 上的场 $\boldsymbol{w}'$，使得在对应点处的向量等长．同样，h' 把 H 上的场 $\boldsymbol{v}$ 投射成 D 上的场 $\boldsymbol{w}$.

一旦完成这证明，$\boldsymbol{v}$ 在 H' 上无零向量的事实将蕴涵 $\boldsymbol{w}'$ 在 D 上无零向量，因而由定理 31.1，指标 $I(w', C) = 0$．然后，作为第二个主要步骤，我们将证明

(2) 在 C 的一点 x 处，向量 $\boldsymbol{w}\boldsymbol{x}$ 与 $\boldsymbol{w}'\boldsymbol{x}$ 是把 $\boldsymbol{v}\boldsymbol{x}$ 绕 C 的切线，先按一个方式旋转 90° 而后按另一个方式旋转 90° 所得到的．这一事实以及 $I(w', C) = 0$ 就蕴涵 $I(w, C) = 2$.

一旦证明了这个，应用定理 31.1 以及 $I(w,C)\neq 0$，就可得到 w 在 D 的某点处有一零向量的结论；从而 v 在 H 的对应点处有一零向量. 所以留待证明的是(1) 与(2).

对于 S 的各点 x，令 T_x 表示过 x 而与 S 相切的平面. 当 x 在 H' 中时，用平行于过 p 与 x 的直线的一族平行线的平行射影来定义映射 $h_x: T_x \to P$. 明确地说，对于 $q\in T_x$，过 q 作一直线平行于过 p 与 x 的直线，则 $h_x(q)$ 是所作的直线与 P 的交点(图 34.2). 证明从 p 到 x 的直线与 T_x 的夹角等于它与 P 的夹角，只是初等几何的一个习题. 所以 h_x 是等距(保持距离) 映射. 定义 D 中的场 $\boldsymbol{w}'$ 如下：$x\in H'$ 给出 D 中的 $hx=y$. P 中 y 处的向量 $\boldsymbol{w'y}$ 定义为 T_x 中向量 $\boldsymbol{vx}$ 在 P 中的 h_x 象. 因为 h_x 是等距映射，向量 $\boldsymbol{w'y}$ 与 $\boldsymbol{vx}$ 等长. 同样，对于 $x\in H$，我们利用平行于过 p' 及 x 的直线来定义 $h'_x: T_x\to P$，并且在 $y=h'x$ 时，$\boldsymbol{wy}$ 是 $\boldsymbol{vx}$ 的 h'_x 象. 这就定义了 D 上的场 w 与 w'，并且完成了(1) 的证明.

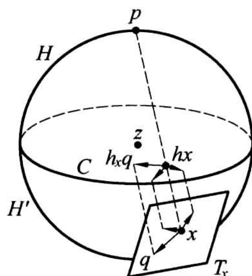


图 34.2

当 $x\in C$ 时，有 $hx=x$ 与 $h'x=x$. 此外，平面 T_x 与 P 都跟过 p 与 x 的直线作成 45° 角. 故从 T_x 映到 P 的映射 h_x 可以看成是把 T_x 绕着 C 在 x 处的切线 L_x 旋转 90° 的结果. 同样地，h'_x 是把 T_x 绕着 L_x 但在相反方向旋转 90° 的结果. 所以作平面 P 关于 C 在 x 处的切线 L_x 的镜面反射，向量 $\boldsymbol{wx}$ 就是 $\boldsymbol{w'x}$ 的反射象.

如同第 30 节中所指出的，D 上的场 $\boldsymbol{w}'$ 可解释为一映射 $f: D\to P$. 因为 $\boldsymbol{w}'$ 无零向量，fD 不包含原点. 把 C 在 D 上缩到 D 的中心点的标准收缩 $\Phi: Q\to D$，与 f 作复合映射 $f\Phi$，就给出从闭曲线 $f\mid C$ 到单独一个点(参看第 24 节与第 25 节) 的一个同伦. 把这同伦的对应于一个值 $\tau\in[0,1]$ 的每一阶段重新解释为 C 上的一个向量场 $\boldsymbol{w}'_\tau$，其中向量都在 P 中. 当 τ 从 0 变到 1 时，我们得到 C 上的一个变动的向量场；那就是，对一固定的 $x\in C$，当 τ 从 0 变到 1 时，向量 $\boldsymbol{w}'_\tau\boldsymbol{x}$ 绕 x 旋转. 当 $\tau=0$ 时，$\boldsymbol{w}'_0$ 场是 $\boldsymbol{w}'$，而当 $\tau=1$ 时，$\boldsymbol{w}'_1$ 是常向量场(全部向量平行并且等长). 因为 fD 不含有原点，向量 $\boldsymbol{w}'_\tau\boldsymbol{x}$ 都不为零.

考虑平面 P 关于 L_x 的反射，定义 $\boldsymbol{w}_\tau\boldsymbol{x}$ 是向量 $\boldsymbol{w}'_\tau\boldsymbol{x}$ 在这反射下的象. 于是，对于各个 τ，$\boldsymbol{w}_\tau$ 是 C 上的一个向量场，其向量都在 P 中. 并且当 τ 从 0 变到 1 时，我们得到从场 $\boldsymbol{w}=\boldsymbol{w}_0$ 到场 $\boldsymbol{w}_1$ 的一个同伦. 因为 $\boldsymbol{w}_\tau\boldsymbol{x}$ 永不为零，从而 $\boldsymbol{w}$ 与 $\boldsymbol{w}_1$ 在 C 上有同一指数. 观察图 34.3，就容易算出指数 $I(\boldsymbol{w}_1,C)$. C 上常向量场 $\boldsymbol{w}'_1$ 用实线向量表示. 场 $\boldsymbol{w}_1$ 用虚线向量表示. 在 C 的各点 x 处，虚线向量 $\boldsymbol{w}_1\boldsymbol{x}$ 是实线向量 $\boldsymbol{w}'_1\boldsymbol{x}$ 关于切线 L_x 的反

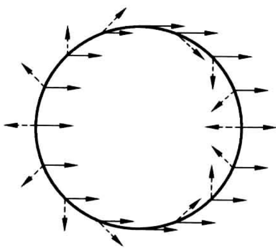


图 34.3

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

射象.在图34.3中,设我们从 C 的最高点开始,沿顺时针方向绕 C 跑一圈.开始 $\boldsymbol{w}_1\boldsymbol{x} = \boldsymbol{w}'_1\boldsymbol{x}$ 指向右.在走过 C 的四分之一时,$\boldsymbol{w}_1\boldsymbol{x}$ 已逆时针方向转了 180°,并指向左.当我们继续绕着 C 跑,它继续以同一速度旋转.当我们绕 $\boldsymbol{C}$ 跑完一圈,它就逆时针方向转了 720°,所以 $\boldsymbol{I}(\boldsymbol{w}_1, C) = 2$.因 $I(\boldsymbol{w}, C) = I(\boldsymbol{w}_1, C)$,从而 $I(\boldsymbol{w}, C) = 2$.这就完成了定理的证明.

35 复 数

一个实变数 x 的某些多项式,例如 $x^4 + x^2 + 1$,没有实零点,这是大家都知道的,这种多项式中最简的是 $x^2 + 1$,它使人们引进了纯虚数 $\sqrt{-1}$,用i表示.然后发现其他多项式的零点都能表成 $a + \mathrm{i}b$ 这种形式,其中 a 与 b 是实数.例如,$x^4 + x^2 + 1$ 的零点是

$$\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}, \frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}$$

$$-\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}, -\frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}$$

这能用直接代入来验证.

我们应该把数系的这种扩大,跟第5节所讨论的各种扩大比较.需要这种扩大的原因完全类似:求解多项式方程时,实数系已显得不够.这就立刻产生两个问题:必须用多大的一类数才能使每个多项式有一个零点?怎样几何地描绘新的数?下一节中我们将要说明:复数集合就足够大了.本节中我们将复习复数的基本性质与它们的几何解释.

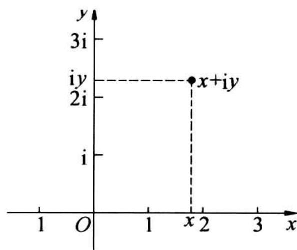


图 35.1

正如实数能用直线上的点表示一样,复数能用平面 P 上的点表示.一个复数 $x + \mathrm{i}y$ 只不过是一对实数 (x, y).在 P 中选定了原点、两垂直的坐标轴和长度单位后,实数对 (x, y) 能画成以 (x, y) 为坐标的点,如图35.1中所表明的.$y = 0$(即形成为 x 或 $(x, 0)$)的复数叫实数.它们是 x 轴上的点.$x = 0$(即 $\mathrm{i}y$ 形式的或 $(0, y)$)的复数叫纯虚数.它们是 y 轴上的点.任何复数 $x + \mathrm{i}y$ 在两坐标轴上的垂直射影是 x 与 $\mathrm{i}y$.实数 x, y 叫做 $x + \mathrm{i}y$ 的实部与虚部.

为着使复数形成数系,必须定义复数的加法与乘法这两种运算.两复数相加是把它们的实部与虚部分别相加

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

或等价地



$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$

加法的几何图形就是向量加法的一几何图形，其中各复数(x,y)表示为从原点到点(x,y)表示为从原点到点(x,y)的向量.两向量的和正好是以这两向量为邻边的平行四边形的对角线(图 35.2).

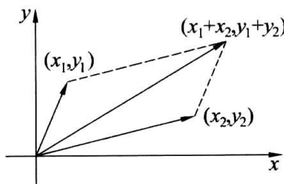


图 35.2

乘法比较复杂.用坐标来表示时，容易用下面的规律来得到乘积

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$$

例如
$$(2, -3) \cdot (\frac{1}{2}, 5) = (14, \frac{23}{2})$$

假定了复数适合分配、结合与交换律，以及 $i^2 = -1$ 这一个额外的规则，就能推导出乘法规则如下

$$(x_1 + iy_1)(x_2 + iy_2) = x_1x_2 + x_1iy_2 + iy_1x_2 + iy_1iy_2 =$$
$$x_1x_2 + i^2y_1y_2 + ix_1y_2 + ix_2y_1 =$$
$$(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

乘法的几何图形根据于向量的夹角以及长度.在原点的角都从正 x 轴量起.于是任何以原点为起点的向量，由一对数$[r,\theta]$决定.其中 θ 是它与 x 轴夹角的度数，$r \geqslant 0$是它的长.两向量(复数)相乘是把它们的角相加，它们的长相乘(图35.3).例如，$i = [1,90°]$，因而 $i^2 = [1,180°] = (-1,0) = -1$.从代数规律推导出这几何规律是三角学中的一个习题.在代数规律中代入 $x_1 = r_1\cos\theta_1, y_1 = r_1\sin\theta_1$ 等，然后应用正弦及余弦的加法公式.

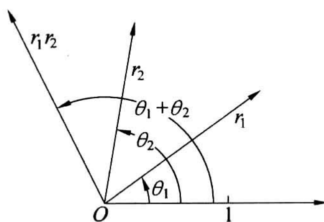


图 35.3

还应做许多工作来说明这些定义合理.第一，对于沿 x 轴的数，必须验证其加法与乘法同实数的一样.这就表明复数形成实数的一种扩展.其次应证明，关于实数的一切代数规律对复数都成立，例如加法与乘法的结合律、交换律以及分配律.实数$1 = (1,0)$是复数的单位，就是$(1,0) \cdot (x,y) = (x,y)$.原点$(0,0)$是加法与乘法两者中的零

$$(0,0) + (x,y) = (x,y), (0,0) \cdot (x,y) = (0,0)$$

最后应该证明加法与乘法都是连续运算.

习惯上总把(x,y)缩写成 z，于是 $z = x + iy$.从而，$z^2 = z \cdot z = (x^2 - y^2, 2xy)$.对所有的整数 $n \geqslant 1$，用归纳法规律 $z^n = z \cdot z^{n-1}$ 来定义 z^n.于是 $fz = z^n$ 定义从复数集合到自身的一个映射 $P \to P$.如果 $n = 1$，f 只不过是恒同映射.

如果 $n = 2$,则 f 是两倍从 x 轴量起的角以及平方从原点量起的距离.从 0 出发的各射线映成两倍角的射线.以 0 为中心、以 r 为半径的圆映射到以 0 为中心、以 r^2 为半径的圆上,并绕它两次.最好把 P 设想为从 0 出发的射线扇形.于是 z^2 覆盖这扇形自身两层.

同样,$fz = z^n$ 把角乘以 n 并且把半径升到 n 次幂.以 0 为中心、以 r 为半径的圆,缠绕以 0 为中心、以 r^n 为半径的圆 n 层.于是,如果 C 是以 0 为中心的任一圆,对于这函数 $fz = z^n$,围绕数 $W(f \mid C,0)$ 是 n.

n 次多项式 f 的定义正如同实数的一样.它是由公式

$$fz = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_2 z^2 + a_1 z + a_0$$

所给出的函数,其中系数 $a_0, a_1, \cdots, a_n$ 都是指定的复数,并且 $a_n \neq 0$.(要记住,每个实数也是一个复数,系数的全部或一部分可以是实数)f 作为一个函数定义一映射 $f: P \to P$.它的连续性可以利用复数的加法与乘法的连续性来证明.

36 每一个多项式都有一个零点

定理 36.1 令 $n \geqslant 1$ 为一整数,又令 f 为以复数为系数的一个 n 次多项式,则 f 至少有一个零点,即至少有一个复数 α 使得 $f(\alpha) = 0$.

因 f 中 z^n 的系数不是零,故能用它去除各项而得 $\frac{f}{a} = g$ 或 $f = ag$,其中 g 的首项系数为 1

$$g(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n$$

因为 g 的一个零点也是 f 的一个零点,只需证明 g 有一个零点.为着这一目的,考虑映射 $g: P \to P$,并将说明有一个圆 C,它的 g 象绕零点 n 次:$W(g \mid C,0) = n$.既然 $n \neq 0$,二维时的主要定理断言 C 内部有一点 α 使得 $g\alpha = 0$.

从 0 到一个复数 z 的距离 $d(z,0)$ 叫做该复数 z 的绝对值,简写为 $|z|$.我们所求的圆 C 将以 0 为中心,而它的半径 r 将是较实数

$$n|a_1|, (n|a_2|)^{\frac{1}{2}}, \cdots, (n|a_n|)^{\frac{1}{n}}$$

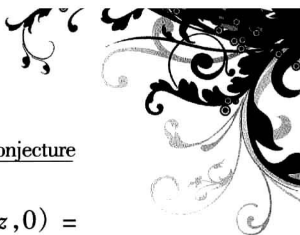
的最大值 r_0 更大的任一数.$W(g \mid C,0)$ 的直接计算太困难,所以我们要建立一个从 $g \mid C$ 到由多项式 z^n 所给定的较简单映射的同伦.在第 35 节中,我们已见到 $W(z^n \mid C,0) = n$.所以如果我们能指出 0 不在这同伦的象中,则围绕数的常值性蕴含 $W(g \mid C,0) = n$.

令
$$h(z) = \frac{a_1}{z} + \frac{a_2}{z^2} + \cdots + \frac{a_n}{z^n}$$

并且用公式

$$g(z,\tau) = z^n[1 + (1-\tau)h(z)], z \in C, 0 \leqslant \tau \leqslant 1$$





来定义 $g \mid C$ 的同伦.当 $\tau = 1$ 时,$g(z,\tau)$ 化简为 z^n;而当 $\tau = 0$ 时,$g(z,0) = gz$.还待证明:0 不在这个同伦的象中,即,对于全体 $z \in C$ 与全体 $\tau \in [0,1]$,$g(z,\tau) \neq 0$.

现在 $z \in C$ 的意思是 $|z| = r$.因 $r > (n|a_k|)^{\frac{1}{k}}$,$k = 1,\cdots,n$,从而 $r^k > n|a_k|$,而这蕴含着对于全体 $z \in C$

$$\frac{|a_k|}{|z^k|} = \frac{|a_k|}{r^k} < \frac{1}{n}$$

因为在 C 上,$h(z)$ 有 n 项,各项的绝对值小于 $\frac{1}{n}$,从而在 C 上,$|h(z)| < 1$.又因为 $|1-\tau| \leqslant 1$,求得 $|(1-\tau)h(z)| < 1$.1 与绝对值比 1 小的复数之和永不为零.所以对于全体 $z \in C$ 与 $\tau \in [0,1]$,$1 + (1-\tau)h(z)$ 不为零.因为当 $z \in C$ 时,z^n 也不为零,从而乘积 $z^n[1+(1-\tau)h(z)] = g(z,\tau)$ 不是零.证明完成.

高于 1 次的多项式仅有一个零点,代数学家对此是不会满意的.下述定理给出完整的结果.

定理 36.2 令 f 为 n 次(实或)复系数多项式,则有 n 个复数 $\alpha_1,\alpha_2,\cdots,\alpha_n$,使得 f 分解成 n 个线性因式的乘积

$$f(z) = a_n(z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_n)$$

如果命 $z = \alpha_i$,第 i 个因子是零;因而每一个 α_i 是 f 的一个零点.如果命 $z = \alpha$,该 α 是不同于 $\alpha_1,\cdots,\alpha_n$ 的数,则每个因子都不是零;从而 $f(\alpha) \neq 0$.这证明了 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 都是 f 的零点,并且它们是仅有的零点.如果一个特别的数在序列中出现两次或更多次,该数就叫做 f 的一个重零点,出现的次数叫做它的重数.

定理的证明需要下述引理.

引理 如果 f 是一个 n 次多项式,α 是任一复数,则有一个 $n-1$ 次多项式 g 使得 $f(z) = (z-\alpha)g(z) + f(\alpha)$.

我们用长除法求 $\frac{f}{z-\alpha}$,来证本引理.令 g 表示商,r 表示余项,有

$$\frac{f(z)}{z-\alpha} = g(z) + \frac{r}{z-\alpha}$$

乘以 $z-\alpha$,得

$$f(z) = (z-\alpha)g(z) + r$$

为了计算 r,命 $z = \alpha$,从而得 $f(\alpha) = r$.

现在应用定理 36.1 来证明定理 36.2,前一定理说 f 至少有一个零点,设为 α_1.因为 $f(\alpha_1) = 0$,引理断言有一个 $n-1$ 次的多项式 g,使得

$$f(z) = (z-\alpha_1)g(z) + f(\alpha_1) = (z-\alpha_1)g(z)$$

因为 $f(\alpha_1) = 0$,所以余项消失了.如果 $n-1 \geqslant 1$,我们可以应用定理 36.1 而获得 g 的一个零点 α_2.于是引理说,有一个 $n-2$ 次的多项式 h,使得

$$g(z) = (z - \alpha_2)h(z)$$

综合这些结果,就给出

$$f(z) = (z - \alpha_1)(z - \alpha_2)h(z)$$

若 $n \geqslant 3$, h 就有一个零点 α_3,并且 $h(z) = (z - \alpha_3)k(z)$;从而

$$f(z) = (z - \alpha_1)(z - \alpha_2)(z - \alpha_3)k(z)$$

每一步把最末尾因式的次数降低 1. 在 n 步之后,末尾因式的次数为零,因而它是常数 c,并且

$$f(z) = (z - \alpha_1)(z - \alpha_2)\cdots(z - \alpha_n)c$$

如果把右端乘出来,z^n 的系数是 c. 所以 $c = a_n$. 证明完成.

历史注释. 定理 36.1 叫做代数的基本定理. 高斯在 1797 年首先给出它的严格证明(见 D. E. Smith, A Source Book of Mathematics, Dover, 1959, 292 页). 在以后的年代中,他给出了几个很不相同的证明,但没有一个像上面所陈述的.

37 结束语:高维情形的一瞥

从一维情形过渡到二维情形时,我们遇到了严重的困难;只能凭借发展一个新概念 —— 闭曲线 φ 在一点 y 处的围绕数 $W(\varphi, y)$,才解决了这困难. 可以预期在过渡到三维以及更高维情形时,会出现另外的困难. 困难确实出现了,不过在二维时已见效的很多想法略加改变都可以推广到高维. 关于这类问题已有很多成果,简短地、粗线条地介绍其中的一部分是值得的. 因为这一部分体现了现代数学研究的一些最好的成果.

在从平面 P 过渡到 n 维欧氏空间 $\mathbf{R}^n$ 时,自然要把圆片与其边界圆代之以 n 维球体 B 与其边界 $n-1$ 维球 S. 令 $f: B \to \mathbf{R}^n$ 为一映射,而 y 是 $\mathbf{R}^n$ 的不在 fS 上的一点. 主要定理说:若 $W(f \mid S, y) \neq 0$,则至少有一个点 $x \in B$ 使得 $fx = y$. 首要问题在于定义 $W(f \mid S, y)$ 这个数,使得它具有 $n = 2$ 时围绕数的所有性质. 当 $n = 3$ 时,最好称 $W(f \mid S, y)$ 为包裹数. 例如,球 S 应该包围 B 的各内点正好一次. 这项工作已完成:对所有维数 n,已精确地定义了数 $W(f \mid S, y)$,并且已证明它具有 $n = 2$ 时的同样性质. 例如,在不通过 y 的任一同伦下,它是不变的.

一旦主要定理已经证明,则在 27 ~ 36 节中所讨论的 $n = 2$ 时的应用,仅需稍许改变符号与说法,就能对所有的维数都可陈述并且证明. 现在让我们陈述其中的几个. 因为球 S 包围球 B 各内点一次,保持 S 的全体点不动的任一映射 $f: B \to \mathbf{R}^n$,具有性质 $fB \supset B$. 其次,任一映射 $B \to B$ 至少有一个不动点. 再者,如果 $\mathbf{R}^n$ 中的一球 S 映到 $\mathbf{R}^{n-1}$,则某些对对径点有同一象点.

有关 $\mathbf{R}^n$ 中球 S 的切向量场的定理必须加以修改. 当 n 是偶数时,S 具有一



连续的非零切向量场(例如,在 $n=2$ 时 S 是一圆). 当 n 是奇数时, S 的任一连续的切向量场至少有一个零.

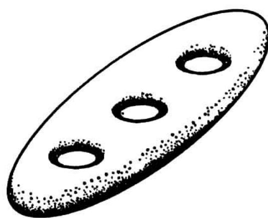


图 37.1

在跳到 n 维情形时,有两个问题我们略去未提,但值得给予更多注意. 其中第一个是:在主要定理中的球体与球面能否用其他图形替代,而不改变其结论的精确性?自然,人们总能用拓扑等价的图形替代,诸如实心盒子以及它的边界面. 但用拓扑相异图形替代时,主要定理的结论能否不引起本质的改变?在维数大于2时答案是"肯定的". 例如,在 $\mathbf{R}^3$ 中令 T 为一环面而 D 为它的内部(例如, D 是一实心环而 T 是它的边界面). 人们能定义一个包裹数 $W(f\mid T,y)$ 使得 T 包裹 D 的各内点一次,而包裹 D 的补集的各内点零次. 在 $\mathbf{R}^3$ 中多连通曲面以及它们的实心内部,提供了 $\mathbf{R}^3$ 中另外的例子. 图 37.1 表示有三个洞的油炸甜饼以及它的边界.

第二个我们曾经略去未提的问题是:为什么我们只考虑 n 维集合映到 n 维集合的映射?我们的主要定理能否推广到容许 k 维集合映到 $\mathbf{R}^n$ 的映射?我们将简略地说明,在 $k=2$ 与 $n=3$ 时,如何能在这方面获得一些成果. 令 $f:D\to\mathbf{R}^3$ 为圆片映到空间的映射. 令 $\varphi:[a,b]\to\mathbf{R}^3$ 为 $\mathbf{R}^3$ 中不与 fC 相交的闭曲线. 对于给定的 $f\mid C$ 与 φ 这两条曲线,规定一个整数 $W(f\mid C,\varphi)$,叫做它们的环绕数. 图 37.2 中的五个例子,按照从左到右的顺序有环绕数 0,1,2,4 与 0. 主要定理的新陈述如下:如果环绕数 $W(f\mid C,\varphi)$ 不是零,则 fD 与闭曲线 φ 至少相交于一点. 读者应把它与图 37.2 中的五个例子相核对,在第一与第五例中,人们能画出一个曲面,它以下面的闭曲线为边界而不与上面的闭曲线相交. 这个曲面会是 fD. 在其他三例中,无论我们如何安装上一个以下面的曲线为边界的曲面 fD,这个曲面总与上面的曲线相交. (在例 3 及 4 中,人们能画出一个扭弯的带子,即一麦比乌斯带,它不与上面的曲线相遇,而它的周界是下面的曲线;但是它不能是一个 fD,因为它是单侧的)

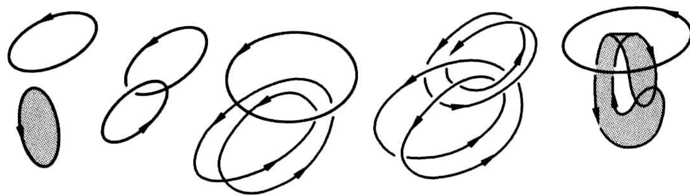


图 37.2

注意:在平面中围绕数 $W(f\mid C,y)$ 涉及一条闭曲线与一个点,而过渡到把一个圆片映到 $\mathbf{R}^3$ 时,替代 y 这个点的是一条闭曲线 φ,围绕数变成环绕数. 如果对于把圆片映到 $\mathbf{R}^4$ 的映射,我们要陈述一个类似主要定理的定理,就会需要 $\mathbf{R}^4$ 中的一条闭曲线 $f\mid C$ 与一个闭曲面 φ 的"环绕数" $W(f\mid C,\varphi)$ 的概念. $\mathbf{R}^4$ 中的

一个闭曲面是一个球面,或一个环面,或多连通闭曲面的任一种映到 R^4 的一个连续映射象.

拓扑学中对于任意维数 n,定义了一个概念叫做圈.一个点,一条闭曲线以及一个闭曲面分别是0维,1维及2维的圈.以圈为边界的图形(例如区间,圆片,球体等)都叫做链.圈,链,它们的同调与同伦,以及它们的交与环绕,形成同调理论的一系列引人入胜题材的主要精华,而同调理论是拓扑学的一个主要部分.

我们已阐明了如何运用拓扑学的一些简单观念来证明定理,这些定理既直觉地使人满意,同时又意义深刻.它们是存在定理.这些定理的高维推广,人们能用同调理论的概念来阐述并且加以证明.





附录 Ⅲ　结构思想十七讲

庞加莱猜想的证明是现代数学的一个重大成就.它综合地应用了现代数学的许多思想和方法.其中颇具法国数学家特点的结构思想.他们认为,全部数学基于三种母结构:代数结构,序结构和拓扑结构.为了更好地理解这一现代数学思想,我们选编了法国索尔本大学教授 G·硕克和法国数学教师联合主席 G·瓦卢辛斯基编的原汁原味的十七个讲座.

1956 年,法国数学会决定与公立学校数学教师联合会一起专为各级学校的数学教师组织几期讲座.

第一期的七个讲座(1956 年 2 月至 6 月)讨论代数结构;第二期的十个讲座(1956 年 11 月至 1957 年 6 月)讨论拓扑结构.

所有的讲座都是在巴黎索尔本大学的昂利－普昂卡雷学院进行的,众多的在职教师专心听讲,新教师与老同行又济济一堂.

这些讲演经各位作者本人撰写成文,分期发表于《数学教师联合会公报》,从而传达给全国各级学校的两千六百多名联合会成员.就我们所知,有些讲稿还曾发表在比利时数学教师联合会的刊物《数学与教育》以及一份意大利的杂志上.

现将这些讲稿汇集在一起,为没有得到这些刊物相应各期(有的已售缺)的读者提供便利;同时,这也体现了教育界的一项合作事业,它的意义是值得强调的.

其实,本文并非研究论著(优秀的论著已经问世).不过是科普作品,因此注定是不痛不痒的.更恰当地说,本文应视为学习和研究当代研究论著的一个入门资料.无论如何,本文将使读者向往了解近代数学各方面的内容.此外,组织这些讲座正是要实现这个目标.讲授已经定型的学科的人与讲授正在成型的学科的人,这两者之间人们感到需要有这样的接触,这一点由于讲座的成功已经得到证明.

研究的领域日益扩展,发现的成果日新月异,这也使从事初等教育的教员们不得不更新他们的理论知识.学院中学到的学问应该定期地加以更新;他们的教学本身应该利用科学的新成就:今日须教今日的数学.这样将使学生对于他们将来的学习或研究获得较好的预备知识.

这些讲座有助于形成一种经常性的教育体系,以满足科学及教育本身的发

展这两方面提出的要求.人们正确地批评了一种只讲“职业诀窍”,“成功之途”而无视科学进步的教育研究.与之相反,这些讲座开辟了俯视当代科学景色的窗口,有可能使教育研究有一个飞跃的进展,适应对科学教育法的大量需要.

瞻望未来,这项研究的前景应该说是乐观的.在这些讲座进行期间,各级现职教师都表现出一种令人十分满意的合作与尊重的气氛.可以一眼看出,各学院的教授们真是尽了最大的努力,对于各位讲演者堪为楷模的全心全意的精神,我们谨致谢忱.同时,如果中学教师能够使他们的学生对于大学学习有较好的预备知识,这将同样是对教授们的帮助.至少,这种合作正好反映了科学教学高度的一致性;科学教学应当采用多种多样适当的方法,使学生一直到大学学习阶段都保持着每个儿童身上几乎总要表现出来的巨大而惊人的创造才能.





第一讲　代数结构[①]

1　引　　言

近几十年来，我们目睹到代数在数学中名副其实地到处渗透.特别是20世纪20年代以后，在E·诺特的推动下，数学家日益清楚地意识到代数的基本概念在数学的几乎所有分支中所起的作用；更确切地说，例如，他们意识到有可能把纯代数中某些多少算是深刻的定理用来考虑分析问题.诚然，这种应用本身并非今日方有（尤其应当提到19世纪末李氏理论的发展）；特别新颖的东西才能引起人们的注意.

随着目前数学的这种代数化，任何研究人员再也不能无视近世代数这一必不可少的工具了.反之，代数也从这种形势下得益不浅，因为拓扑与分析不断向代数提出一些新问题，产生了几十年前也许几乎无法想象的进展.

中学的数学教学，至少在最后一个学年，理应受到这个演变的影响.如果说要重视新观点，问题无疑不在修订教学大纲，而在如何阐述古典理论.这无疑就是法国数学会发起组织这一系列讲座的理由.

然而代数是什么呢?用几句话给出代数的定义，使得代数与其他数学分支的界限一目了然，这是不容易的.诚然，在科学发展的每个时期，每位数学家对什么是代数与什么不是代数的看法是足够清楚的.但是任何明确的定义在以后科学的发展中都有变得陈旧或过于狭隘的危险.事实上，不可能预先给代数划定什么不可逾越的范围（任何科学分支的情形也是如此），因为无法预见到在探索过程中会显现出哪些新领域.

粗略地说，可以认为代数是研究过一个或几个集合的元素施行的某些运算，而不考虑这些元素本身的性质.对于给了某些运算的一个集合，一切所能阐述的内容也完全适用于与它同构的任何其他集合（后文中要介绍同构的概念）.对代数的这种理解可能一个世纪以来都占上风，然而最近的进展无疑必将使代数扩大其过于狭窄的范围，因而上述理解今日可能已经过时了.这里，我们只限于用几个例子来说明我所谓的经典的代数概念，至于开创了目前某些进展的那

① 原作者H.Cartan（索尔本大学教授）.

些新观点,过几十年再让别人向读者说明吧!

2 运算的概念

从算术起就有了运算的概念.运算是一个法则:对于两个元素 a 与 b,相应地给出一个元素 c(元素 c 有时称为 a 与 b 的和,有时称为它们的积,有时还有另外的名称).其实,这里不过就是一个函数的概念:给了三个集合 A,B 与 C,考虑由元素 $a \in A$ 与元素 $b \in B$ 形成的偶对 (a,b) 的集合,叫做 A 与 B 的积集合,记作 $A \times B$;于是,一个运算就是定义在集合 $A \times B$ 上并在集合 C 中取值的函数.我们也说,运算是把 $A \times B$ 映入 C 的一个映射.

一个重要情形是这三个集合 A,B 与 C 相同的情形,此时,考虑的是一个映射 $f: A \times A \to A$,这样的函数叫做内合成法则.如果只假定 $B = C$,则得到所谓的外合成法则:即是把 $A \times B$ 映入 B 的函数.此时,每个元素 $a \in A$ 确定一个把 B 映入 B 的映射,即是与 b 相应的是 $f(a,b)$.不过,把 B 映入 B 的映射也称作 B 的一个变换,于是外合成法则相应于 A 的每个元素给出 B 的一个变换;集合 A 就叫做算子域,并说 A 作用于 B.

下面举几个合成法则的例子.整数加法:与一对整数对应的是一个整数.整数乘法也一样.这些都是整数集合中的内合成法则.分数、实数、复数的加法与乘法,以及一个变量或多个变量的多项式的加法与乘法,也都是内合成法则.在初等几何中有位移概念.我们知道相继施行位移 a 与位移 b 就得到一个位移 c,有时称为 a 与 b 的积(或合成).这样,相应于一对位移 (a,b) 我们得到一个位移 c,这是位移集合中的内合成法则.

下面是另外一些例子.设 X 与 Y 是同一集合 E 的两个子集,其交 $X \cap Y$ 是 E 中同时属于 X 与 Y 的那些元素组成,其并 $X \cup Y$ 是 E 中至少属于集合 X 与 Y 之一的那些元素组成.在 E 的所有子集构成的集合 A 中,对于偶对 (X,Y) 相应地给出交 $X \cap Y$ 的法则是一个内合成法则;同样,对于 (X,Y) 相应地给出并 $X \cup Y$ 的法则也是 A 中的内合成法则.

按照另一种思想,我们也会考虑两个整数 a 与 b 以及它们的最大公约数 $d(a,b)$;映射 $(a,b) \to d(a,b)$ 是整数集中的内合成法则.最小公倍数情形也是一样.

再给出字的例子.设有集合 E,所谓"字"是指 E 中元素组成的有限序列 uvwxy,这些元素可以不同,也可以相同.其中还有一个"空字",即是空序列定义的字.给了两个字 uvwxy 与 pqrs,相继写下这两个字可确定出一个新字 uvwxypqrs.这是 E 的元素生成的字集 A 中的一个内合成法则.如果将字 uvw 与空字合成,仍得到同一个字 uvw.





上述所有例子都是内合成法则的例子.现在举两个外合成法则的例子.考虑初等几何中的平面(或空间),取定一点 O.对每个实数 t,我们使以 O 为中心,以 t 为比例常数的位似与之相应;这是一个变换.这里我们得到一个外合成法则,其算子域 A 是实数 t 的集合,相应的集合 B 是几何学中的平面(或空间).

再考虑空间的"图形"(例如三角形).一个位移将图形变成图形.于是若取位移集合为 A,空间图形的集合为 B,就得到一个外合成法则,其算子域是位移集合.

以后我们几乎只讨论内合成法则.

3 内合成法则的各种性质

综合性 一个合成法则,比如记作乘法(ab 表示 a 与 b 的合成),称为结合的,如果对任意的 a,b 与 c,有 $(ab)c = a(bc)$.

在上面例子中我们给出的所有内合成法则都是结合的.容易给出非结合法则的例子:对一对实数 (a,b),我们使其和之半 $\frac{a+b}{2}$ 与之相应;立即可以验明这个法则不是结合的.

如果一个法则是结合的,可以定义任何有限多个元素的序列 $a_1,\cdots,a_n$ 的合成,并证明有关结合性的一个一般定理;例如,有

$$(ab)(cdef)g = (abc)(de)(fg)$$

交换性 一个内法则(为确定计记作乘法)称为交换的,如果对任意 a 与 b,有 $ab = ba$.在上述例子中,实数、复数以及多项式的加法与乘法都是交换法则.同样,集合的交运算与并运算也是交换的;求两个整数的最大公约数或最小公倍数的法则也是交换的.与此相反,几何中位移的合成不是交换的;"字"的合成也不是交换的.

这样,就存在结合的但不是交换的法则.同样也有一些交换的但不是结合的法则,例如求两实数 a 与 b 之和之半的法则.

一个法则如果既交换又结合,则可以定义任意元素组(自然是有限组——译注)的合成,而不必计及它们的次序.

中性元 给了一个内法则(为确定计记作乘法),元素 e 称为中性元,若对每个元素 a,有 $ae = ea = a$.这种元素不一定存在,但若存在必唯一,因为若 e 与 e' 是两个中性元,则有 $e = ee' = e'$.

逆元 考虑一个有中性元 e 的内法则.我们说 a 与 b 互为逆元,若 $ab = ba = e$.

如果 a 至少有一个逆元,并且法则是结合的,那么 a 的逆元是唯一的;因为,若 b 与 b' 都是 a 的逆元,则有

$$b = be = b(ab') = (ba)b' = eb' = b$$

如果考虑的法则表作加法(通常只是在法则是交换的时候才行),其中性元若存在,一般记作0(零).我们有

$$a + 0 = 0 + a = a$$

此时,“逆元”改称为“反元”:若 $a + b = 0$,则称 a 与 b 互为反元.

4 由一个或多个合成法则定义的代数结构

设 E 是一集合.给了一个或多个合成法则,就在 E 上定义了一个代数结构.所有的代数结构(不论定义在哪个集合上),只要定义它们的法则数目一样,满足同样一些明确罗列的条件,就都用同一个名称.用几个例子来说明这点.

群 集合 E 上的群构造是由一个合成法则定义的,这个法则必须满足下列条件(或“公理”):

公理1 法则是结合的;

公理2 存在中性元;

公理3 每个元素具有逆元.

具有这样一个合成法则的集合 E 就称为群.

例如,整数(正、负及零)的加法是一个群法则.实数或复数的加法也是群法则;具有前述合成法则的位移集合是一个群.非零实数或非零复数的乘法是一个群法则.反之,正整数的乘法法则不是群法则,因为一个正整数一般没有逆元.集合的交运算法则或并运算法则也都不是群法则.同样,“字”的合成法则不是群法则.

一个群称为交换群(或阿贝尔群),若其合成法则是交换的.

环 一个集合 A 上的环构造是由两个合成法则定义的,分别称为加法与乘法;这两个法则必须满足下列公理:

公理1 加法是一个阿贝尔群法则(其中性元记作0).

公理2 乘法是结合,有中性元(称做单位元),记为1;约定 $1 \neq 0$.

公理3 乘法对于加法是分配的

$$a(b + c) = (ab) + (ac),(b + c)a = (ba) + (ca)$$

易见,每个元素与0的积等于0;因而0没有乘法逆元.一个环称为交换环,若其乘法是交换的.

环的例子:全体整数;实数或复数集;实系数或复系数的多项式.(其加法与乘法运算如平常所定义)

体 一个体是一个环 A 并满足下面补充条件:每个非零元素都有乘法逆元素;这也就是说,A 的非零元素构成一个乘法群.





例子 整数集不是一个体,多项式全体不构成一个体;有理数集,实数集以及复数集都是体.

当然,存在各种各样的群,环与体.伽罗瓦已经确定了所有的有限体,亦即具有有限多个元素的体.最简单的是只有两个元素0与1的体,其加法与乘法是显然的(特别有 $1+1=0$).对每个素数 p 与每个整数指数 $f\geqslant 1$,存在恰好有 p^f 个元素的体,并且这样的体是唯一的.记整数 p^f 为 n,体中每个元素 x 满足方程 $x^n-x=0$(这就给出一个例子,说明存在系数不全为零的多项式,对变量所有的值均为零;这个多项式是 n 次的,并且变量仅能取 n 个不同的值,因为体中仅有 n 个元素).如果整数 n 不是 p^f 形状的数,则不存在由 n 个元素数组成的体.

5 同构概念

方才说存在唯一一个具有 p^f 个元素的体.这不完全对.正确的说法是:如果两个体有同样多个元素,则它们是同构的.就是说在这两个体的元素之间存在一个一一对应,并且这个对应保持加法与乘法.一般,对各有其代数结构的两个集合 E 与 E',可以建立同构的概念;例如,两个群之间的同构是这两个群的元素之间的一一对应,并保持每个群的合成法则;例如,指数函数 $f(x)=a^x$(这里 a 是不等于1的正常数)就是把实数加法群映成正实数乘法群的同构,因为 $f(x+y)=f(x)f(y)$.又两个环 A 与 A' 之间也有同构概念:即是一个一一对应,保持加法与乘法,还使 A 与 A' 的单位元彼此对应.

代数对象实际上只研究到同构为止.这正是实质所在,因为我们只关心所研究运算的性质.如果有两位数学家谈到整数环,我们并不问他们是否谈的是同样的对象,因为从他们对整数环研究出来的那些性质推知,任何两个环如果具有这些性质就必然同构.

6 从已有的代数对象构造新的代数对象:多项式环的例子

设给了一个交换环 A(例如实数体或整数环,等).我们来定义一个新环,称为字母 X 的(形式)多项式环.它的元素用行号 $\sum_{n\geqslant 0}a_nX^n$ 表示,这里诸 a_n 是 A 中元素,只有有限多个不为零;记号 X^n 暂时没有任何内容,只是一个形式记法,因此,定义多项式的,不过就是一序列"系数" $a_0,a_1,\cdots,a_n,\cdots$,罢了,这些"系数"除有限多个外全为零.在多项式集合中定义加法如下:多项式 $\sum_{n\geqslant 0}a_nX^n$ 与多项式 $\sum_{n\geqslant 0}b_nX^n$ 的"和"定义为多项式 $\sum_{n\geqslant 0}(a_n+b_n)X^n$,其系数是相应系数之和 a_n+

b_n. 立刻可见这个加法是交换的，结合的，具有零元 0(所有系数全为零的多项式)，并且每个多项式都具有"反元"；换句话说，多项式集构成加法群.

p 次单项式是指一个多项式 $\sum_{n\geqslant 0} a_nX^n$，使得 $n \neq p$ 时所有系数 a_n 皆为零；这样的单项式可以简记作 a_pX^p. 我们看到每个多项式 $\sum_{n\geqslant 0} a_nX^n$ 是其各次单项式 a_nX^n 的和. 这就反过来说明了所用记法的理由.

多项式的乘法定义如下：首先定义单项式 a_pX^p 与单项式 b_qX^q 的积为单项式 $(a_pb_q)X^{p+q}$，其次数 $p+q$ 是两个单项式的次数之和，其系数 a_pb_q 是系数 a_p 与 b_q 的积. 为了定义两个多项式 $\sum a_nX^n$ 与 $\sum b_nX^n$ 的积，先把每个多项式分解为它的不同次数的单项式，然后对这些单项式两两求积再相加；这给出下列公式：积是多项式 $\sum c_nX^n$，这里 $c_n = \sum_{p=0}^{n} a_pb_{n-p}$.

可以验明，字母 X 的多项式集 $A[X]$，照上面那样定义加法与乘法后，成为一个交换环，其单位元是系为 1 的零次单项式. 这样就定义了字母 X 的多项式环(系数在交换环 A 中).

X 纯粹是一个符号. 但我们可以用"值"来替代它. 给了一个元素 $x \in A$，约定：对每个单项式 a_nX^n，我们相应地有 A 的元素 a_nx^n(系数 a_n 与 x 的 n 次幂的积)；然后，对多项式 $\sum a_nX^n$ 也相应地有这些 a_nx^n 的和. 这样就得到 A 的一个元素，称为这个多项式在 X 取值 x 时的"值". 当 x 在 A 中变化时，一个多项式的"值"就是 x 的一个函数，其值属于 A. 这样，每个多项式都定义一个函数，但一开始就应该提防将形式多项式与它所确定的函数混为一谈.

现取定元素 $\alpha \in A$，对每个多项式 $\sum a_nX^n$，我们有与它相应的值 $\sum a_n\alpha^n \in A$. 这就得到一个映射 f，把环 $A[x]$ 映入环 A；这个映射使得对应于两个多项式之和的值，就是这两个多项式之值的和，而对应于两个多项式之积的值，则是这两个多项式之值的积. 我们定义形式多项式的加法与乘法的方式，正是为了保证映射 f 具有上述性质. 于是，设 u 与 v 为两个多项式，则有

$$f(u+v) = f(u) + f(v), f(uv) = f(u)f(v), f(1) = 1$$

只要两个环 B 和 A 以及把 B 映入 A 的映射 f 满足这些条件，就说 f 是把 B 映入 A 的一个同态. 同态的概念可以推广至其他种种代数结构中；粗略地说，同态是一个保持合成法则的映射.

再来讨论多项式，我们可以将前面所述推广. 设 A' 是包含 A 的一个环；可以将字母 X 换成属于 A'(不再只属于 A)的值 x；这时，每个多项式 $\sum a_nX^n$ 取值 $\sum a_nx^n$，这个值是 A' 的元素. 只要 x 与 A 的元素可交换，即对于每个 $a \in A$，有 $ax = xa$，上述做法仍然定义了一个把环 $A[x]$ 映入环 A' 的同态. 例子：设 A 是





实数体,E 是三维空间中由原点 O 发出的向量构成的向量空间. 把 E 映入 E 的线性映射构成一个环(我们知道如何定义两个线性映射的和,相继施行两个线性映射就得到它们的积);这个环 A' 包含了以0为中心的全体位似构成的环 A,环 A 可以与纯量体(即实数体)等同. 于是,如果在实系数多项式 $\sum a_nX^n$ 中,将 X 换成把 E 映入 E 的线性映射 x,就给出了 $\sum a_nx^n$,这仍然是一个把 E 映入 E 的线性映射. 这个法则使得与每个多项式对应的是一个线性映射,这对研究向量空间的线性变换十分重要. 这个例子使我们看到从形式多项式概念得到的好处,即是可以将"变量"X 换成系数环的元素以外的东西. 另外一个例子是把 X 换成线性微分方程理论中讲的那种微分算子.

7 商结构

从已有代数结构造出新的代数结构时(例如,从整数环造出有理数体或从实数体造出复数体),常常需要等价关系的概念. 集合 E 的元素 x 与 y 之间的关系 $R(x,y)$ 称为一个等价关系,如果

(1)$R(x,x)$ 总是真的("反身律");

(2)$R(x,y)$ 蕴涵 $R(y,x)$("对称律");

(3)$R(x,y)$ 与 $R(y,z)$ 蕴涵 $R(x,z)$("传递律").

下面是代数中出现的等价关系的若干例子:

1) 设 m 为已知整数,两个整数 x 与 y 之差 $x-y$"被 m 整除"这个关系是整数 x 与 y 之间的等价关系.

2) 更一般地,设 G 为乘法群(元素 x 的逆元记作 x^{-1}),H 为 G 的子群,即是含单位元 e 的一个子集,使得 $x\in H$,与 $y\in H$ 蕴涵 $x^{-1}y\in H$. 对于 G 的元素 x 与 y,关系 $x^{-1}y\in H$ 是 G 中的一个等价关系. 例1)的情形就是:G 是整数加法群,H 是 m 的倍数构成的子群(这时用加号表示合成法则).

3) 设 (a,b),(a',b') 是两对整数,满足 $b\neq 0$,$b'\neq 0$. 若 $ab'=a'b$,则称它们等价. 需要验明,这的确是一个等价关系. 易见这个关系是反身的与对称的,于是只要说明它是传递的即可. 这有三对整数 (a,b),(a',b') 与 (a'',b''),满足 $ab'=a'b$,$a'b''=a''b'$,问题是要证明 $ab''=a''b$. 将式子 $ab'=a'b$ 两边乘以 b'',注意乘法的结合性与交换性,有 $ab'b''=a'bb''=ba'b''$,据 $a'b''=a''b'$,前式后一元素等于 $ba''b'$,于是 $ab'b''=ba''b'$,从而 $b'(ab''-a''b)=0$;因为 $b'\neq 0$,而一个乘积只有当一个因子为零时才能为零,所以得到 $ab''=a''b$.

注意,上面的论证对于下面更一般场合仍然有效. 不考虑整数对,而考虑一个环 A 的元素对 (a,b),这里 $b\neq 0$,这时要求 A 满足下面条件:A 是交换环,且当 u 与 v 不为零时,积 uv 不能为零. 这种环称为整环.

整环的例子：设 A 是交换环，$A[x]$ 是字母 X 的多项式环，P 与 Q 为两个非零的多项式，亦即其系数不全为零，所以有 $P = a_pX^p + a_{p+1}X^{p+1} + \cdots, a_p \neq 0$ 与 $Q = b_qX^q + b_{q+1}X^{q+1} + \cdots, b_q \neq 0$. 于是积 PQ 是单项式 $(a_pb_q)X^{p+q}$ 及次数 $> p+q$ 的单项式的和. 如果系数环 A 是整环(例如 A 是体)，则积 $a_pb_q \neq 0$，从而乘积多项式 PQ 不为零. 这就证明了环 $A[X]$ 也是整环. 这时我们就可以在多项式对 (P,Q) 与 (P',Q')（这里 $Q \neq 0, Q' \neq 0$）之间定义一个等价关系，这实质上就是有理分式的等价，这里无须诉诸函数的概念而是纯粹形式地把有理分式的等价化成多项式的恒等.

现在再来讨论等价关系的一般理论. 设 R 是集合 E 中的等价关系，x 是 E 的元素，所有与 x 等价的元素的集合称为 x 的等价类. 条件(1)，(2) 与(3) 说明 x 属于 x 的等价类；若 x 与 y 是 E 的两个元素，则 x 的等价类与 y 的等价类或者完全相同或者没有公共元素. 所有等价类的集合记作 E/R，也称为 E 对等价关系 R 的商集合. 我们也说 E/R 是把 E 中关于 R 等价的元素黏合而得.

现在再考虑上面讨论过的那些等价关系的例子1)，2) 与3). 例1) 的等价类称为模 m 整数(或模 m 等价类)；商集合称为模 m 整数集. 例2) 的等价类称为 G 对子群 H 的一个右陪集；$x \in G$ 的右陪集记作 xH，由形如 xu 的元素组成，这里 u 遍取 H 中元素. 例3) 的等价类正是一个分数(当 A 是多项式环时，就是“有理分式”).

8 与等价关系相容的合成法则

设 R 是集合 E 中的等价关系，还给了 E 中一个内合成法则(为确定计记为乘法). 如果合成元 xy 的等价类仅仅依赖于 x 的等价类与 y 的等价类，就得到等价类集合 E/R 中的一个合成法则.(此时亦说这个合成法则与等价关系相容)

例子 设 x 与 y 为两个整数，考虑它们的和 $x+y$. $x+y$ 的模 m 类并不因用 x' 代替 x 而有所改变，只要 $x-x'$ 为 m 整除；它也不因用 y' 代替 y 而有所改变，只要 $y-y'$ 为 m 整除. 由此可得模 m 整数的加法，这是模 m 整数集中一个合成法则；这是交换的与结合的. 同样，积 xy 的等价类仅仅依赖于 x 与 y 的等价类；由此可得模 m 整数的乘法. 容易验明，模 m 整数全体构成一个交换环. 此外，若 m 是素数，它还是一个体.

现在定义分式的乘法①：在偶对 $(a,b)(b \neq 0)$ 集合中考虑合成法则：$(a,b)(a',b') = (aa',bb')$，这是交换的与结合的法则，有单位元 $(1,1)$. 可以证明，(aa',bb') 的等价类仅仅依赖于 (a,b) 与 (a',b') 的等价类；由此可得两

① 这时相应的环 A 自然假设是整环. —— 译者注





个分式的积.于是分式的乘法是交换的与结合的;此外,每个不在(0,1)的等价类中的元素(正确地说,应是这个元素所属的等价类)具有逆元,因为当 $a \neq 0$, $b \neq 0$ 时,(a,b) 与 (b,a) 的积等价于 $(1,1)$.

注 利用偶对 (a,b) 的集合中下述合成法则,容易定义分式的加法,这个法则只是当 $b = b'$ 时才对 (a,b) 和 (a',b') 有定义,即是 $(a,b) + (a',b') = (a+a',b)$.将此法则过渡到商集后,环 A 的分式集合 Q 就有了加法与乘法,而且成为一个环,也是一个体.此外,对每个元素 $a \in A$,使一个分式,即 $(a,1)$ 的等价类与之对应,我们看到,这就得到一个同构 f,把环 A 映成体 Q 的一个子环.这时我们说,f 可以把 A 粘到体 Q 的一个子环上.(例如,可以把整数环粘到有理数体的一个子环上,也可以把字母 X 的多项式环粘到字母 X 的有理分式体的一个子环上)

最后例子 拓扑空间 X 的基本群(亦称"普昂卡雷群").

我们不来定义所谓的拓扑空间,而只指出若干的实例:通常的欧氏空间、球面、环面以及带孔的圆盘,等.

让我们确切地定义空间 X 中的道路,即曲线段的概念:这就是指把数直线的线段 $[0,1]$ 映入空间 X 的一个连续映射 $t \to f(t)$.我们指出,对于"参数" t 的不同的值,$f(t)$ 可取相同的值;这点是没有妨碍的.

此后取定空间 X 的一点 M.以 M 为起点的闭路定义为一条道路 $t \to f(t)$,使得 $f(0) = M$ 与 $f(1) = M$.以 M 为起点的闭路集合 L 中,定义一个合成法则如下:给了闭路 f 与 g,定义闭路 h 为

$$h(t) = f(2t) \quad (0 \leqslant t \leqslant \frac{1}{2})$$

$$h(t) = g(2t-1) \quad (\frac{1}{2} \leqslant t \leqslant 1)$$

换句话说,在头一半"时间"里跑过闭路 f,然后在后一半时间里跑过闭路 g.这个法则不是结合的,因为如果有三条闭路 f,g 与 h,那么闭路 $(fg)h$ 是头四分之一的时间跑过 f,第二个四分之一的时间跑过 g,而在后一半时间里跑过 h;然而,闭路 $f(gh)$ 则是在前一半时间里跑过 f,第三个四分之一时间里跑过 g,而最后四分之一时间里跑过 h.

现在,在闭路集合 L 中引进一个等价关系 R:闭路 f 称为等价于(或说同伦于)闭路 g,如果存在一个连续形变,即是连续变化的一族以 M 为起点的闭路,把 f 变成 g;精确地说,这是指存在一个连续函数 $F(t,u)$,两个变量 t 与 u 在 $[0,1]$ 中变化,使得

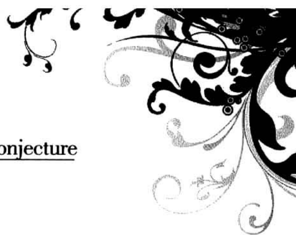
$$F(t,0) = f(t), F(t,1) = g(t)$$
$$F(0,u) = F(1,u) = M$$

可以验明,这的确是一个等价关系,并且闭路的合成法则与这个关系相容;这

样,在闭路等价类的集合 L/R 中就有一个合成法则,这一次是结合的了.可以证明这是一个群法则,其单位元 e 是满足下述条件的闭路 f 的等价类:当 $0 \leqslant t \leqslant 1$ 时,$f(t) = M$(即是坚守在点 M 不动的闭路).设有闭路 f,由 $f'(t) = f(1-t)$ 定义的闭路 f'(这是同一条闭路,但以“相反的方向”跑过)具有下述性质:f 和 f' 的等价类在闭路等价类的群 L/R 中是互逆的.

方才定义的群(空间 X 以 M 为基点的基本群)在拓扑学中与在函数(自守函数)论中都起重要的作用.从最后这个例子我们看出代数的概念如何进入数学的其他领域.





第二讲　环,同余,理想[①]

1　定义与若干简单性质的回顾

考虑任一集合 $A = \{a, b, \cdots\}$,集 A 称为一个环,如果 A 上定义了两个运算,亦即内合成法则,分别叫做加法与乘法,用通常的记号表示,满足下述公理:

Ⅰ. A 关于加法是阿贝尔群,即加法是结合的:$a + (b + c) = (a + b) + c$;交换的:$a + b = b + a$(“阿贝尔”这个词就表示这个交换性质);存在单位元,叫做零元,记为 0,满足 $a + 0 = a$;对每一元 $a \in A$ 有一反元 $-a$ 满足 $a + (-a) = 0$(简写成 $a - a = 0$).

Ⅱ.乘法是结合的:$a(bc) = (ab)c$.

Ⅲ.乘法关于加法是分配的

$$a(b + c) = ab + ac, (b + c)a = ba + ca$$

上述各种性质当然是假定对 A 中任何 a, b, c 都成立.

若乘法是交换的,环就叫可换的.本文讨论的主要是可换环.

在定义环的公理中,没有假设存在单位元 e 或 1(乘法单位元:对任意的 $a \in A$,有 $ea = ae = a$)[②].如果存在单位元,就称 A 是有单位元的环.在只由一个零元素构成的环内,零就是乘法单位元,不过,我们提到有单位元的环时,一般不是指这种情形.

易见,在每个环 A 内,乘法关于减法是分配的,零元素 0 满足:$0 \cdot a = a \cdot 0 = 0$(对任意 $a \in A$),即只要有一因子为零乘积就为零,但这乘积为零的充分条件只有对一些特殊的环才是必要条件;我们要碰见一些很简单的环,有些非零元之积可以为零.

多于一个元素的环 $K, K \neq \{0\}$,关于乘法决不能成群,这是因为

$$a \neq 0, a \cdot 0 = 0 \cdot a = 0$$

但是,$K - \{0\}$ 关于 K 的乘法有可能成群(即 K 中两个非零元的积不是零,存在单位元 e,并且每一元 $a \neq 0$ 具有逆元素 a^{-1}:$aa^{-1} = a^{-1}a = e$).如果这

① 原作者 P. Dubreil(索尔本大学教授).

② 在这点上,我们的定义与 H · 嘉当在其关于代数结构的讲演中所给出的定义不同,要一般些,而且是沿用经典的定义.

样,就说 K 是体.

2 例

全体整数的集合 $\mathbf{Z}=\{\cdots,-m,\cdots,-1,0,1,2,\cdots,n,\cdots\}$ 是环;全体偶数的集合也是环,因为我们不要求环具有单位元.

有理数集 $\mathbf{Q}$,实数集 $\mathbf{R}$,复数集 $\mathbf{C}$ 都是体(包含集合 $\mathbf{Z}$),代数数的集合 Σ 也是体.所谓代数数,就是有理系数或整系数(两者是一回事)代数方程的根(实数或复数)①.

给了一个环 A,例如上面诸例之一,我们可以定义系数在 A 中的字母 X 的多项式环 $A[X]$②,以及系数在 $A[X]$ 中的字母 Y 的多项式环 $A[X][Y]$;后面这种多项式可以写成 $\sum_{ij}a_{ij}X^iY^j$, $a_{ij}\in A$(诸系数 a_{ij} 中仅有有限多个不为零),关于字母 X 和 Y 显然对称,这就得到所谓的系数在 A 中的字母 X 与 Y 的多项式环,记成 $A[X,Y]$.可以递推定义系数在 A 中的,n 个未定元 $X_1,X_2,\cdots,X_n$ 的多项式环 $A[X_1,\cdots,X_n]$.

另一个重要的环是模 m 整数环,m 是给定的整数.这个环是用模 m 的算术同余关系 R 定义的,其中加法与乘法与同余关系是相容的(也可以说,这个同余关系关于整数环 $\mathbf{Z}$ 的加法与乘法是一个正规等价关系):这个环就 $\mathbf{Z}$ 关于(同余)等价的商集 $\mathbf{Z}/R$ 或 $\mathbf{Z}/(m)$,可以表示成(以 m 作除数所得 m 个余数的)集合 $\{\bar{0},\bar{1},\cdots,\overline{m-1}\}$.对于每个给定的 m,容易造出模 m 整数集 $\mathbf{Z}/(m)$ 的加法与乘法表,从而可以发现一个值得注意的情况:例如,对于 $m=4$ 我们有 $\bar{2}\cdot\bar{2}=\bar{0}$;对于 $m=6$,我们有 $\bar{2}\cdot\bar{3}=\bar{0}$,一般,只要 m 不是素数,$\mathbf{Z}/(m)$ 就不是整域(或整环);反之,若 p 为素数,则 $\mathbf{Z}/(p)$ 是整域,甚至是体(稍过片刻我们就要定义整域).

在复数体 $\mathbf{C}$ 中,我们考虑如 $\alpha+\beta\mathrm{i}$ 的复数,这里 α,β 是有理数:$\alpha,\beta\in\mathbf{Q}$.这些数的集合显然是体,记为 $\mathbf{Q}(\mathrm{i})$.记号 $\mathbf{Q}(\mathrm{i})$ 表明,这里 $\mathbf{C}$ 中包含体 $\mathbf{Q}$ 与数 i 的最小子体.我们说,$\mathbf{Q}(\mathrm{i})$ 是由添加 i 而得.可以不使用 $\mathbf{C}$,而从 $\mathbf{Q}$ 出发用下述方法直接构造体 $\mathbf{Q}(\mathrm{i})$,这是后面要讨论的符号添加法的特例.

在多项式环 $\mathbf{Q}[x]$ 中,考虑多项式 x^2+1,现在让它担任前面讨论整数环 $\mathbf{Z}$ 时模 m 的那种角色.我们说,$\mathbf{Q}[x]$ 的两个多项式 $\varphi(x)$ 和 $\psi(x)$ 有关系 R,如果 $\varphi-\psi$ 是 x^2+1 的倍数.这样定义的关系 R 是一等价关系,并且关于多项式的加

① 这个命题容易归结为对称函数的理论.

② 参看 H. Cartan 关于代数结构的讲演.字母 X 是"纯符号",也称为未定元,其作用犹如循环半群的生成元.在有单位元的环中,X 是一个特殊的多项式($a_1=e$;当 $i\neq 1$ 时 $a_i=0$).





法与乘法是正规的.等价的多项式构成的类,其相加与相乘如同所含多项式的相加与相乘,于是多项式等价类的集合构成环,称为商环,或剩余类环,记为 $\mathbf{Q}[x]/R$ 或 $\mathbf{Q}[x]/(x^2+1)$. $\mathbf{Q}$ 中不同的数属于不同的等价类,可以把两者看成一样.每个等价类中都含有一个一阶多项式 $\alpha x+\beta$,即是类中各个多项式用 x^2+1 去除所得的公共余式.若用 i 表示包含多项式 x 的等价类,则包含 $\alpha x+\beta$ 的等价类是 $\alpha\mathrm{i}+\beta$(根据正则性).因为 x^2+1 与 0 是两个等价的多项式,故有 $\mathrm{i}^2+1=0$.最后,若 α,β 不全为零,则 $\alpha+\beta\mathrm{i}(\neq 0)$ 有逆元 $\dfrac{\alpha-\beta\mathrm{i}}{\alpha^2+\beta^2}$.这个商环正好就是包含 $\mathbf{Q}$ 与方程 $x^2+1=0$ 的根 i 的体(显然,不存在任何更小的体具有此性质).

上面的纯代数方法起始于柯西,如用实数体 $\mathbf{R}$ 代替有理数体 $\mathbf{Q}$,也可用这种方法造出复数体 $\mathbf{C}=\mathbf{R}(\mathrm{i})$.我们也可以把 -1 换成一个不是完全平方的整数 d,仍按上述方法构造出体 $\mathbf{Q}(\sqrt{d})=\mathbf{Q}[x]/x^2-d$,特别,可以用纯代数的方法界定形如 $\sqrt{d}$ 的无理数.上面的方法叫做符号添加法,可以用来定义任意的代数数,这是后话.今后要着重考虑关于环的加法与乘法的正规等价关系;为简化说法以及联想到这一概念的算术来源,我们把这些关系称为同余关系.

在如上造出的体 $\mathbf{Q}(\sqrt{d})$ 中,有一个简单但值得注意的环:就是形如 $a+b\sqrt{d}$ 的数所成的环,这里 a,b 是任何整数: $a,b\in B$.特别,数 $a+b\mathrm{i}$ 是高斯整数;高斯整数构成的环由平面的"整点格"表示,所谓整点就是坐标为整数的点.

虽然我们只讨论交换环,但要提醒,一个环 A(可以是体)中的元素组成的 n 阶方阵构成一个环($n>1$ 时为非交换环).

3 整 域

整域(或整环)是一个交换环 D,使得 D 内的乘积 xy 仅当因子 x 或 y 至少有一为零时才能为零.换句说话: $D-\{0\}$ 中两个元素的积仍在 $D-\{0\}$ 中,即 $D-\{0\}$ 是乘法封闭的.此外, $ab=ac$ 与 $a\neq 0$ 蕴涵 $b=c$;形如 $ab=ac$ 的等式在 $D-\{0\}$ 可约简.若 $D-\{0\}$ 为有限集,则在其乘法表中,一个元在每一行或每一列正好出现一次,这就是说,方程 $ax=b$ 恒有解,故 $D-\{0\}$ 是乘法群(当然是阿贝尔群),于是 D 是体;每一有限整域都是体. p 为素数时, $\mathbf{Z}/(p)$ 就是一例.

若 D 是非有限的整域,考虑分式 (a,b) 或 $\dfrac{a}{b}(a,b\in D,b\neq 0)$,便可造出

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

一个含有 D 的(最小的)体,称为整域 D 的商体.这个体的元素是分式的等价类①.

设 K 是(交换)体;若 $x,y\in K,xy=0,y\neq 0$ 时,x 在 K 内有逆元素 x^{-1},于是

$$0=x^{-1}(xy)=(x^{-1}x)y=y$$

所以,K 中任何一个环,特别是 K 自身,都是整域.

今考虑多项式环 $D[x]$,这里 D 是整域,n 次多项式 $f(X)=a_0+a_1X+\cdots+a_nX^n(a_n\neq 0)$ 与 p 次多项式 $g(X)=b_0+b_1X+\cdots+b_pX^p(b_p\neq 0)$ 的积是 $n+p$ 次多项式 $a_0b_0+\cdots+a_nb_px^{n+p}$(因为 a_nb_p 不为零).若 A 是任意环,那么,在环 $A[x]$ 内,n 次多项式与 p 次多项式的积至多是 $n+p$ 次;若 A 不是整域,则存在乘积的次数小于 $n+p$ 的例子.当 A 不是整域时,有个很值得注意的意外:一多项式 $f(x)\in A[x]$.当 $x=a,x=b$ 时为零,$a,b\in A,b\neq a$,我们并不能断言此多项式能被 $(x-a)(x-b)$ 整除;此外,一个 n 次多项式 $f(x)$ 完全可能对于 n 个以上不同的 x 值为零.例如,取 $A=\mathbf{Z}/(4)$,一次多项式 $2x$ 当 $x=0$ 与 $x=2$ 时为零.

4 特征数

设环 A 具有单位元 e,n 是整数.若 n 是正整数,$x\in A$,则令

$$nx=x+\cdots+x\quad(n\text{ 个相加})$$

若 $n=0(\in\mathbf{Z})$,我们令 $0x=0(\in A)$.若 n 是负整数,我们令

$$nx=-(|n|x)$$

因 $x=ex$,故有

$$nx=(ne)x$$

因此,若存在非零整数 n,使得

$$ne=0\qquad(*)$$

则对所有 $x\in A$ 有

$$nx=0$$

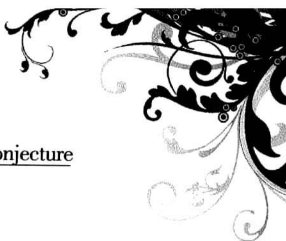
设 n_0 是满足如上等式的正整数中之最小者,这时,我们说,n_0 是环 A 的特征数.

若不存在任何非零整数满足方程(*),就说 A 的特征数为零(因为这时 0

---

① 参看 H. Cartan 关于代数结构的讲演.D 中单位元的存在不是必须的(我们用分数 $\frac{m}{m}$ 代替 $\frac{1}{1}$,用 $\frac{am}{m}$ 代替 $\frac{a}{1}$).构造一个整域的商体,可以更简单地从构造包含已知阿贝尔半群 S 的最小群而得(集合 S 是阿贝尔半群是指:S 具有结合的、交换的乘法运算,并满足约简规则:$ab=ac$ 蕴涵 $b=c$).阿贝尔半群的结构也完全可以用来从自然数出发定义整数,或者从正实数出发定义任意实数,这里,所考虑的运算是加法.





是满足下述条件的唯一整数:对任何 $x \in A$ 有 $0e = 0, 0x = 0$).

整域 D 的特征数或为 0 或为一素数 p. 否则,若 $n_0 = n'_0 n''_0, n'_0, n''_0 \neq 0, 1$,则 $n'_0 e = a \neq 0, n''_0 e = b \neq 0, ab = n_0 e = 0$,这样 D 就不是整域了.

请特别注意,教学工作中经常有人说,"所有代数计算规则" 都是特征为零的整域内有效的计算规则,我想这是不确切的.

现在让我们来考察特征数 $p \neq 0$ 的整域内某些有趣的特点.

在每个交换环内,二项式公式成立

$$(a + b)^n = a^n + C_n^1 a^{n-1} b + \cdots + C_n^k a^{n-k} b^k + \cdots + b^n$$

$$C_n^k = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}$$

(在非交换环内,例如我们只能写出 $(a + b)^2 = (a + b)(a + b) = a^2 + ab + ba + b^2$)

在上面二项式公式中,取 n = 整域 D 的特征数 p,因 $k \neq 0, k \neq p$ 时,C_p^k 是 p 的倍数,故有

$$(a + b)^p = a^p + b^p$$

用 $a - b$ 代 a 得

$$(a - b)^p = a^p - b^p$$

逐步使用这公式,又得

$$(a_1 + \cdots + a_r)^p = a_1^p + \cdots + a_r^p$$

$$(a + b)^{p^f} = a^{p^f} + b^{p^f}$$

$$(a - b)^{p^f} = a^{p^f} - b^{p^f} (f \text{ 是任意自然数})$$

设 D 具有单位元 e. 在上面公式中,取 $a_1 = a_2 = \cdots = a_r = e$,将 re 写成 r,则对任意的自然数 r 有

$$r^p = r$$

特别是,这个公式对模 p 整数体 $\mathbf{Z}/(p)$ 成立. 如果仍然用整数环 $\mathbf{Z}$ 的语言,可见对所有自然数 x 有费马定理

$$x^p \equiv x(p)$$

现在设 A 是任意交换环,在多项式环 $A[x]$ 中,可用纯代数方法定义多项式 $f(x)$ 的导数:把 $f(x + h)$ 按 h 的幂展开

$$f(x + h) = f(x) + hf'(x) + \cdots \quad f(x + h) \in A[x, h]$$

其中 h 的系数就定义为 $f(x)$ 的导数. 由此容易得到导数的计算法则:若

$$f(x) = a_n x^n + \cdots + a_1 x + a_0$$

则

$$f'(x) = n a_n x^{n-1} + \cdots + a_1$$

$f'(x)$ 至多是 $n - 1$ 次多项式;当环 A 的特征数是 0 时,$f'(x)$ 确为 $n - 1$ 次;反之,若 A 的特征数是 n_0,而 n 是 n_0 的倍数,则 $f'(x)$ 的次数肯定小于 $n - 1$. 特别,

若 A 是具有单位元的整域或特征数为 p 的体,则 x^p 的任何多项式

$$f(x) = a_n x^{pr} + \cdots + a_1 x^p + a_0$$

有零导数(即"恒为零").可以想见,这结果在代数方程的理论中是轰动一时的,若取 A 是特征数为 p 的某些体,则上述结果与 $f(x)$ 在 $A[x]$ 中的既约性并非不相容.例如①A 是未定元 t 的,系数在体 $K = \mathbf{Z}/(p)$ 中的有理分式体 $K(t)$;$f(x) = x^p - t$.

5 理　想

在交换环 A 中,设 R 是同余关系(即对 A 的加法与乘法的正规等价关系).等价类的集合或商集 A/R 是一个环,简单一些就记成 $\overline{A}$.设 $\overline{x}$ 是包含 A 的元 x 的等价类.映射 $x \to \overline{x}$ 是把 A 映成 $\overline{A}$ 的映射($\overline{A}$ 的每个元 $\overline{x}$ 至少是 A 的一个元的象),并且这映射还保持 A 与 $\overline{A}$ 的加法与乘法,即

$$\overline{x + y} = \overline{x} + \overline{y}, \overline{xy} = \overline{x}\,\overline{y}$$

具有上述性质的任何映射 $A \to \overline{A}$ 称为(环的)同态映射;环 $\overline{A}$ 叫做与 A 同态,也叫做 A 的同态象,记为

$$A \sim \overline{A}$$

因此,A 内每一同余关系 R 都产生一个与 A 同态的环 $\overline{A}$,也就是商环 A/R.

考察 A 的元素 0 所属的等价类 $\overline{0}$,$\overline{0}$ 显然是环 A/R 的零元,但同时又是 A 的一个子集合 M,具有如下两个重要性质:

1° M 是 A 的加法子群;事实上,由定义 M 包含 0,又 $x, y \in M$ 蕴涵 $x + y$,$-x \in M$.(利用一个群的子群 M 的简单特性,可以更简捷地验明 $x, y \in M$ 蕴涵 $x - y \in M$)

2° A 对 M 可以施行乘法,即对任意的 $a \in A$,$m \in M$ 必有 $am \in M$,这是因为 $am \equiv a0 = 0(R)$.

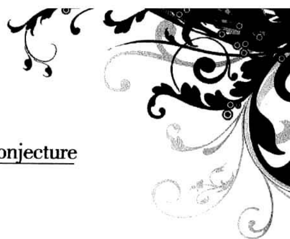
因为 R 关于加法是正规的,所以 $x \equiv y(R)$ 与 $x - y \equiv 0(R)$,即 $x - y \in M$ 同时成立.若已知零类 M,由此可得同余关系 R.

交换环 A 的一个子集如果具有上述性质 1° 与 2°,则称为理想;按照上面的论证,一个同余关系 R 的零类就是理想.反之,假定先有 A 的一理想 M;如果 A 的两个元素 x 与 y 满足 $x - y \in M$.就定义 x 与 y 有关系 R,则 R 显然是一等价关系,并且关于 A 的加法与乘法是正规的,所以 R 是一同余关系,其零类恰好是 M.

这样,交换环 A 的所有理想 M 与所有同余关系 R 成一一对应.因此,商环

① 这个例子的讨论要用到几个眼下无法介绍的命题.





A/R 可写成A/M.

现在假定先有环的同态映射

$$A \to \overline{A}$$

在这映射下,A 的元x 以$\overline{A}$ 的元$\overline{x}$ 作为象.对于 A 的两个元素x 与y,我们说 x 与 y 有关系R,如果 $\overline{x} = \overline{y}$.容易验明,$R$ 是一个等价关系(称为同态等价),并且是一同余关系.于是可以构造出商环 A/R.按作法,关于 R 的等价类,即是 A/R 的元素,与$\overline{A}$ 的元成一一对应.因为这个一一对应保持加法与乘法,所以是同构映射.因此,$\overline{A}$ 与商环$A/R = A/M$ 同构.理想 M 或模R 零类是A 的这样一些元素的集合,它们被这个同态映射映成 $\overline{A}$ 的零元素(我们说 M 是这个同态映射的核).

所以,在同构的意义下,A 的同余关系,即是 A 的理想就确定了所有与A 同态的环.

6 特殊情形与例子

在交换环 A 的理想所成的集合中,定义如下运算:交(在集合论的意义下),和以及积.A 的两个理想 $\mathfrak{a}$ 与 $\mathfrak{b}$ 的和 $\mathfrak{a}+\mathfrak{b}$ 是形如 $a+b$ 的元素所成的集合,这里 $a \in \mathfrak{a}, b \in \mathfrak{b}$.这显然是包含 $\mathfrak{a}$ 与 $\mathfrak{b}$ 的理想.积 $\mathfrak{ab}$ 是由积 ab 生成的理想,即和 $a_1b_1 + \cdots + a_rb_r$(r 不固定)的集合.这里 $a_i \in \mathfrak{a}, b_i \in \mathfrak{b}$.为避免累赘,关于理想的运算,我们不多介绍,其实它们是很重要的,而且可以据以定义商以及剩余.

现在让我们只讨论一种基本的生成方式;为简便计,假定环 A 具有单位元 e.

在环 A 中取定 h 个元素 $a_1, \cdots, a_h$,容易验明,形如 $a = x_1a_1 + \cdots + x_ha_h$ $(x_i \in A)$ 的元素的集合是一个理想 $\mathfrak{a}$,记为

$$\mathfrak{a} = (a_1, \cdots, a_h)$$

我们说,$\mathfrak{a}$ 是由 a_i 生成的,$a_1, \cdots, a_h$ 是它的(有限)基.由一个元素生成的任何理想 $\mathfrak{a} = (a)$ 叫做主理想.单位元 e 生成的主理想(e) 就是环 A,这个理想常叫单位理想.元素 0 生成的主理想就是 0,称为零理想.

在体 K 内,只有零理想与单位理想.事实上,设 M 是K 的理想;$M \neq 0$.于是 M 至少含有一个元 $a \neq 0$;因 K 是体,故这个元素 a 有逆元素 a^{-1},因之 $e = a^{-1}a \in M$,故 $K = (e)$ 含于 M,由此 $M = K = (e)$.

在整数环 $\mathbf{Z}$ 内,以及在系数属于体 k 的 x 的多项式环$k[x]$ 内,所有的理想都是主理想.

就环 $k[x]$ 为例证明上述结论.零理想当然是主理想.今设 $\mathfrak{a} \neq (0)$ 是一理

想，f 是 $\mathfrak{a}$ 中一个最低次的非零多项式，g 是 $\mathfrak{a}$ 中任一多项式. 用 f 除 g：$g = qf + r$，这里，r 的次数低于 f 的次数. 多项式 r 属于理想 $\mathfrak{a}$，故必为零多项式，所以 g 属于主理想 (f)，从而 $(f) = \mathfrak{a}$.

对于整数环 $\mathbf{Z}$，证明类似：在这个环内，除模 m 确定的算术同余外，不存在别的同余关系；m 是与算术同余相应的主理想的生成元.

一个环，如果其中存在带剩余的除法，就叫欧氏环；一个环的每个理想如果都是主理想，就叫主理想环（简称主环），前面的论证不难推广：每个欧氏环都是主理想环. 高斯整数环也是欧氏环（为定义带剩余的除法，利用元素 $\alpha = a + \mathrm{i}b$ 的范数 $N(\alpha)$，而范数 $N(\alpha)$ 的定义是 $N(\alpha) = a^2 + b^2$，在未定元的多项式除法中起次数的作用）.

多个未定元的多项式环 A，例如 $\mathbf{Q}[x,y]$，$\mathbf{C}[x,y]$，不再是主理想环：事实上，理想 (x,y) 不可能有唯一的生成元；注意，这个理想有一个简单的几何意义：它是 A 中在原点为零的多项式的集合. 设 K 是任一体，n 元多项式环 $K[x_1,\cdots,x_n]$ 中，理想的生成遵从一个简单的法则：在这样的环中，每个理想都有有限基. 具有这种性质的环，称为诺特环（诺特是德国女数学家 Emmy Noether 的姓. 她利用“极大条件”与“因子链的公理”给出了另一些公理定义，并得到诺特环中理想理论的基本结果）.

刚才提到的“环 $K[x_1,\cdots,x_n]$ 是诺特环”的定理是希尔伯特得到的. 确切点说，这是一类“迁移定理”的一种表现形式：具有单位元的环 A，如果是诺特环，则 $A[x]$ 也是诺特环.

7 一些重要的理想

交换环 A 的诸理想 M 中，使得相应的商环即是剩余类环 A/M 具有重要性质的理想，自然更应重视. 让我们来研究使 A/M 是体或整域的理想 M.

理想 M 称为无因子理想或极大理想，如果 A 是唯一的严格包含 M 的理想. 我们有下述定理：

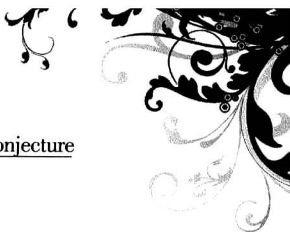
欲使 A/M 是体，M 必须是无因子理想. 若 A 具有单位元，这必要条件还是充分的.

证明 1° 设 A/M 是体.

令 M' 是严格包含 M 的理想：$M \subset M'$，设 $a \in M'$，$a \notin M$（不属于 M）. 用 $\bar{a}$ 表示包含 a 的等价类：$\bar{a}$ 不是零类 $M = \bar{0} \in A/M$. 因此，若 x 是 A 的任意元素，$\bar{x}$ 是含有 x 的等价类. 则存在 $\bar{y}(= \overline{a^{-1}x})$，使得 $\bar{x} = \bar{a}\bar{y}$，从而若 y 表示 A 中属于等价类 $\bar{y}$ 的一个元素，则

$$x \equiv ay(M)$$





换言之

$$x = ay + m \quad m \in M \subset M'$$

从 $a \in M'$ 得 $x \in M'$,故 $A = M'$;M 是无因子理想.

2° 设 A 具有单位元 e,M 是无因子理想.为了证明 A/M 是体,只需证明每个不是 $\bar{0} = M$ 的等价类 $\bar{a}$ 都有逆元.属于类 $\bar{a}$ 的元素 a 不属于 M:理想 $M + (a)$ 严格包含 M,由此

$$M + (a) = A$$

特别,单位元 e 也属于左端,即存在 $m \in M$,$r \in A$,使得

$$e = m + ra$$

由此,取模 m 等价类,便得

$$\bar{r}\bar{a} = \bar{e}, \text{即} \bar{r} = \bar{a}^{-1}$$

上述定理有如下的重要应用.考虑体 K 以及环 $K[x] = A$ 的既约多项式 $f(x)$,即 $f(x)$ 不是同一环内次数起码为 1 的两个多项式的积.A 中的主理想 (f) 是无因子理想,故 $A/(f)$ 是体,记为 K'.因 K 的两个相异元不能模 f 同余,故含有 K 中元素的等价类与 K 的元成一一对应;这些等价类构成一个与 K 同构的体,将这个体与 K 等同,于是 K 是 K' 的子体.

若 α 是含有多项式 x 的等价类,则 $f(\alpha)$ 是含有多项式 $f(x)$ 的等价类(根据同态映射 $A \sim A/(f)$);但含有 $f(x)$ 的等价类是 K' 的零元素.故在 K' 内有 $f(\alpha) = 0$,这样,我们就构造出一个体 K',在其中 $f(x) = 0$ 至少有一个根:本质上讲,这就是符号添加法.

环 A 的理想 $\mathfrak{p}$.如果使得环 $A/\mathfrak{p}$ 是整域,就叫做素理想.$\mathfrak{p}$ 为素理想的充要条件是:在 $A/\mathfrak{p} = \bar{A}$ 内下列关系成立

$$\bar{x} \neq \bar{0}, \bar{y} \neq \bar{0} \text{ 蕴涵 } \bar{x}\bar{y} \neq \bar{0}$$

所以在 A 内

$$x \notin \mathfrak{p}, y \notin \mathfrak{p}, \text{蕴涵 } xy \notin \mathfrak{p}$$

这表明补集合 $A - \mathfrak{p}$ 是乘法封闭的.

素理想这个概念在代数理论本身以及算术、几何的应用中,都是极重要的.

我们只指出,在有单位元的环 A 内,每个无因子理想 M 都是素理想(事实上,A/M 是体,因而是整域).

最后,作为例子,考虑复系数三元多项式环 $A = \mathbf{C}[x,y,z]$,以及理想

$$\mathfrak{a}_1 = (x,y,z), \mathfrak{a}_2 = (x,y), \mathfrak{a}_3 = (x)$$

$A/\mathfrak{a}_1$ 与 $\mathbf{C}$ 同构,所以是体;$A/\mathfrak{a}_2$ 与多项式环 $\mathbf{C}[z]$ 同构,$\mathbf{C}[z]$ 是整域;$A/\mathfrak{a}_3$ 与多项式环 $\mathbf{C}[y,z]$ 同构,也是整域.诸理想 $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3$ 满足下列包含关系

$$\mathfrak{a}_3 \subset \mathfrak{a}_2 \subset \mathfrak{a}_1 \subset A = (1)$$

每一个都是素理想,$\mathfrak{a}_1$ 还是无因子理想.

第三讲　向量空间,线性形式与线性方程①

1　向量空间

中等教育的许多材料已涉及向量的概念:平移的合成、速度的合成、力的合成;纯量积,向量积,一个向量关于一点的矩,再往后,就是一些隐约可见的代数形式,例如多项式、线性形式等.因之,使用向量这个字眼,其含意并不总是一样的:时而是约束向量,时而是滑动向量,时而又是自由向量.

本讲只讨论自由向量.

我们常常利用向量的几何表示,即是一端带有箭头的直线段,这可能使我们以为,可以孤立地考虑一个向量,而不管别的向量如何,尽管如此,但不久我们就会看到,要把一个数学对象叫做向量,它必须和另一些相似的对象共同组成一个集合,可以对这个集合施行某些有精确定义的运算.

这些对象之所以能够叫做向量,正是由于这些运算的本性,而不是每个对象的个别特性.例如多项式,连续函数,方阵可以叫做向量,正是这个道理.

含混地说,数学对象构成的一个集合可以叫做向量的集合,如果我们能在该集合中把两个向量相加,同时又可用任何一个数去乘每个向量.

精确一点,我们给出下面的定义.

定义　实数体 $\mathbf{R}$ 上的向量空间是指一个集合 E,满足下述条件:

1) E 上定义了一个内合成法则②,叫做加法,E 关于加法成为交换群.

2) E 上定义了一个外合成法则:对于每个数 $\lambda,\mu\in\mathbf{R}$ 与每个 $\boldsymbol{x}\in E$,相应地给出 E 的一个元 $\lambda\boldsymbol{x}$,使得:

a) $\lambda(\boldsymbol{x}+\boldsymbol{y})=\lambda\boldsymbol{x}+\lambda\boldsymbol{y}$;

b) $(\lambda+\mu)\boldsymbol{x}=\lambda\boldsymbol{x}+\mu\boldsymbol{x}$;

c) $\lambda(\mu\boldsymbol{x})=(\lambda\mu)\boldsymbol{x}$;

d) $1\cdot\boldsymbol{x}=\boldsymbol{x}$.

集合 E 的元叫向量;$\mathbf{R}$ 的元叫纯量.

---

① 原作者 G. Choquet(索尔本大学教授).

② 参见 H. Cartan 关于代数结构的讲演.由于 E 是交换群,所以用 $x+y$ 表示 x 与 y 的合成,$-x$ 表示 x 的反元.



R 的 0 与群 E 的零元素，显然并非同一元素，但我们同时用 0 表示.

从上面的关系 a),b),c),d) 容易得出如下关系

$$0\boldsymbol{x} = 0, \lambda 0 = 0, (-\lambda)\boldsymbol{x} = \lambda(-\boldsymbol{x}) = -\lambda\boldsymbol{x}$$

关系 a),b),c),d) 只涉及 **R** 上存在加法与乘法；因此，似乎也可定义环 A 上的向量空间，这个环 A 起前面定义中 **R** 的作用. 不过，只有当环 A 是一个体时，才能得到与 **R** 上的向量空间类似的理论. 这主要是因为，A 是一个体时，关系 $\lambda\boldsymbol{x} = 0$ 蕴涵 $\lambda = 0$ 或 $\boldsymbol{x} = \mathbf{0}$，但若 A 仅为一个环，上述蕴涵关系未必定成.

最常用的体有实数体 **R** 与复数体 **C**.

上面向量空间的定义中，没有用到体 **R** 具有序结构这一事实. 只有当人们希望在(**R** 上的) 向量空间 E 中定义半直线以及区间的概念，从而可以谈凸集时，才涉及序结构. 正是这个事实，才使 **R** 上的向量空间具有一种特殊的作用. 本文不打算引入半直线与凸集的概念.

向量空间的例子

1) n 个实数所成序列 $\boldsymbol{x} = (x_1, \cdots, x_n)$ 的集合是 **R** 上的向量空间，只要令

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n)$$

$$\lambda\boldsymbol{x} = (\lambda x_1, \lambda x_2, \cdots, \lambda x_n)$$

这个空间就是欧氏空间 $\mathbf{R}^n$.

特别是 $n = 1$ 时，**R** 自身可看成 **R** 上的向量空间.

2) 定义在区间[0,1] 上的数值函数的集合是 **R** 上的向量空间，只要定义 $h = f + g$ 与 $k = \lambda f$ 为：对每个 $x \in [0,1]$

$$h(x) = f(x) + g(x), k(x) = \lambda f(x)$$

这个空间的下列子集合也是 **R** 上的向量空间：[0,1] 上的连续函数的集合. [0,1] 上的无限可微函数的集合；多项式(或者 n 次多项式) 的集合；三角多项式的集合，等.

3) 复平面 **C** 上的正则解析函数的集合是体 **C** 上的向量空间.

要确认一个集合是否为向量空间，并非都是一样容易的事. 例如，考虑 [0,1] 上的连续函数的集合，其中每个函数的表示曲线均有有限长度. 两个这样的函数的和是否仍有同样的性质，并不显然. 又如，令 E 是实数的无限序列 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n, \cdots)$ 所成的集合，每个序列都满足

$$x_1^2 + x_2^2 + \cdots + x_n^2 + \cdots < +\infty$$

要确认两个这种序列的和 $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \cdots, x_n + y_n, \cdots)$ 仍有同样的性质，也不显然.

同构的概念

设 E 是次数 $< n$ 的实系数多项式所成的向量空间. 若 $p = (a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n)$ 是 E 的元，注意 p 由其系数所成的序列$(a_1, \cdots, a_n)$ 决定，由

b) E 的每个自由子集至多含有 n 个元;若恰好含有 n 个元,它就是一个基底.

c) E 的每个向量子空间 V 至多是 n 维;若它是 n 维,就是 E 本身.

d) E 的每个异于 E 的向量子空间都有互补的子空间,这两个互补子空间维数的和等于 E 的维数 n.

3 线性映射

在古典分析与现代分析中,占主导地位的概念之一是线性函数的概念,这个概念只是在向量空间的范围内才表现出它的全部意义.

定义 若 E 与 F 是在同一体 K 上的两个向量空间.把 E 映入 F 的线性映射,是指把 E 映入 F 的任何映射 f,使得对任何 $\boldsymbol{x},\boldsymbol{y}\in E,\lambda\in K$,有

$$f(\boldsymbol{x}+\boldsymbol{y})=f(\boldsymbol{x})+f(\boldsymbol{y}),f(\lambda\boldsymbol{x})=\lambda f(\boldsymbol{x})$$

例如,把 E 映成 F 的同构映射不外乎就是把 E 映成 F 的一一线性映射.

线性映射的例子

1) E 是向量空间 $E_1,E_2,\cdots,E_n$ 的积.对于 E 中任何 $\boldsymbol{x}=(\boldsymbol{x}_i)$,把 E 映成 E_i 的映射 $\boldsymbol{x}\to\boldsymbol{x}_i$ 是线性的.

2) 若 X_1 与 X_2 是空间 E 的两个互补子空间,把 E 映成 X_1 的映射 $\boldsymbol{x}\to pr_1\boldsymbol{x}$ 是线性映射.

3) 令 P_n 是一个变元的 n 次多项式 $p(x)$ 构成的向量空间.若 $f(p)$ 表示 p 的 n 次项,则把 P_n 映入 P_n 的映射 f 是线性映射.

4) 令 E 是 $[0,1]$ 上的可微实函数 $x(t)$ 构成的空间,用 x' 表示函数 x 的导数:

a) 把 E 映入 $\mathbf{R}$ 的映射 $x\to x'(t)$ 是线性映射.

b) 把 E 映入 $[0,1]$ 上的数值函数构成的空间的映射 $x\to x'$ 是线性映射.

5) 令 E 是 $[0,1]$ 上的实连续函数空间:

a) 定义 $X(t)=\int_0^t x(u)\mathrm{d}u$,每个 x 都有其相应的 X,映射 $x\to X$ 是把 E 映入 E 的线性映射.

b) 把 E 映入 $\mathbf{R}$ 的映射 $x\to\int_0^1 x(u)\mathrm{d}u$ 是线性映射.

6) 令 E 是定义在圆盘 $x^2+y^2\leqslant 1$ 上的连续函数空间;F 是定义在该圆周界上的连续函数空间.考虑 E 的元素 φ 在圆周上的限制,这个对应关系是把 E 映入 F 的线性映射.

一些明显的性质

其中最重要者列出如下:





向量子空间·一个子集生成的子空间

在 $\mathbf{R}^3$ 中,非 0 的向量子空间是通过原点的诸直线与平面.

在一个变元的多项式空间中,n 次多项式的集合,可以被一个已知多项式整除的多项式的集合,以及只有偶次项的多项式的集合,都是向量子空间.一般有如下定义.

定义 设 E 是向量空间.如果 E 的子集 A 满足:$\boldsymbol{x}$ 与 $\boldsymbol{y} \in A$ 蕴涵 $\boldsymbol{x}+\boldsymbol{y} \in A$ 以及对任意纯量 λ,$\lambda \boldsymbol{x} \in A$,就称 A 是 E 的向量子空间.

显然,E 的每个向量子空间都是 E 的加法子群.

另一方面,若 $(A_i)_{i\in I}$ 表示 E 的一族向量子空间,则属于每个 A_i 的元素的集合,即诸 A_i 的交集合,仍为 E 的向量子空间.

特别,命 X 为 E 的任一子集.E 中包含 X 的向量子空间,除 E 外可能还有别的;所有这些向量子空间的交集 $A(X)$ 显然是 E 中包含 X 的最小向量子空间,我们说 $A(X)$ 是 X 生成的向量子空间.

现在引入一个新概念来确定 X 的元素的性质:

定义 命 X 是向量空间 E 的子集合.E 中每个形如 $\boldsymbol{x} = \sum\limits_{i\in I} \lambda_i \boldsymbol{x}_i$ 的元称为 X 的元素的线性组合,这里 $(\boldsymbol{x}_i)_{i\in I}$ 是 X 中一族有限多个元素,而诸 λ_i 是纯量.

X 的元素的任何线性组合都属于包含 X 的任何向量子空间,特别是属于 $A(X)$;另一方面,所有的线性组合显然构成 E 的一个向量子空间;这个子空间包含 X,所以也包含 $A(X)$,它又含于 $A(X)$,因此两者恒同.换言之:

E 的子集 X 生成的子空间 $A(X)$ 恒同于 X 的元素的线性组合构成的集合.

互补子空间

在 $\mathbf{R}^3$ 中,平面与从原点出发且与平面只在原点相交的直线是互补子空间的例子.

定义 X_1 与 X_2 是 E 的两个子空间,如果 E 的每个点 $\boldsymbol{x}$ 都能唯一地写成

$$\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 \quad \boldsymbol{x}_1 \in X_1, \boldsymbol{x}_2 \in X_2$$

就说 X_1 与 X_2 互补.

容易验明,上述条件相当于 $(X_1 \cup X_2)$ 生成 E 且 $X_1 \cap X_2 = \{0\}$.

我们说,$\boldsymbol{x}_1$ 是 $\boldsymbol{x}$ 在 X_1 中的分量,或者说 $\boldsymbol{x}_1$ 是 $\boldsymbol{x}$ 在 X_1 上平行于 X_2 的投影.

若记 $\boldsymbol{x}_1 = pr_1 \boldsymbol{x}$,则可验证

$$pr_1(\boldsymbol{x}+\boldsymbol{y}) = pr_1\boldsymbol{x} + pr_1\boldsymbol{y}, pr_1(\lambda\boldsymbol{x}) = \lambda pr_1\boldsymbol{x}$$

例子 1) E 是 x 的实系数多项式集合;X_1 是能被 x^2+1 整除的多项式集合;X_2 是次数不高于 1 的多项式集合.

2) E 是 $\mathbf{R}$ 上的函数空间;X_1 是偶函数 $(f(-x) = f(x))$ 构成的子空间;X_2 是奇函数 $(f(-x) = -f(x))$ 构成的子空间.

在空间 $\mathbf{R}^3$ 中,一个平面的互补子空间全都是直线,就是说,互补子空间总

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

是同构的,这是如下一般命题的特例:

命题 在向量空间 E 中,若 X_2 与 X'_2 是同一子空间 X_1 的两个互补子空间,则平行于 X_1 的投影是 X_2 与 X'_2 之间的同构映射.

证明 对每个 $\boldsymbol{x}'_2\in X'_2$,$\boldsymbol{x}'_2=\boldsymbol{x}_2+\boldsymbol{x}_1$.这里 $\boldsymbol{x}_2$ 是 $\boldsymbol{x}'_2$ 在 X_2 上平行于 X_1 的投影.但这个等式也可写在 $\boldsymbol{x}_2=\boldsymbol{x}'_2+(-\boldsymbol{x}_1)$,故 $\boldsymbol{x}'_2$ 又是 $\boldsymbol{x}_2$ 在 X'_2 上平行于 X_1 的投影.

因此,如上的(投影)映射 $\boldsymbol{x}_2\to\boldsymbol{x}'_2$ 是把 X_2 映成 X'_2 的一一映射;因为是投影,也有

$$(\boldsymbol{x}_2+\boldsymbol{y}_2)'=\boldsymbol{x}'_2+\boldsymbol{y}'_2,(\lambda\boldsymbol{x}_2)'=\lambda\boldsymbol{x}'_2$$

故这是同构映射.

2 线性无关·基底

几何中,讨论把 $\mathbf{R}^3$ 的一个向量沿着三个已知向量的方向进行分解时,已碰见过线性无关的概念;在代数中,研究线性方程组的可解条件时也碰见这个概念.我们要说明,这个概念怎样才能不依赖于所研究向量的性质而且与坐标系的特殊选取无关.

定义 向量空间 E 的一族有限多个元素 $(\boldsymbol{a}_i)_{i\in I}$ 叫做自由的,如果每一关系 $\sum\limits_{i\in I}\lambda_i\boldsymbol{a}_i=0$ 蕴涵对每个 i 有 $\lambda_i=0$.

如果一族任意多个向量的每一个有限子族都是自由的,就说该向量族是自由的.

非自由的族叫做约束族.

若一个族是约束族,则存在形如 $\sum\limits_{i\in I}\lambda_i\boldsymbol{a}_i=0$ 的一个关系,其中有一个 $\lambda_i\neq 0$.例如 $\lambda_0\neq 0$ 时,可得 $\boldsymbol{a}_0=\sum\limits_{i\in J-0}\mu_i\boldsymbol{a}_i$.反之,只要这样的关系成立该族必是约束的.故:

一个族是约束族等价于其中有一个元素是其余元素的线性组合.

自由族的诸元称为线性无关的;约束族的诸元则称为线性相关的.

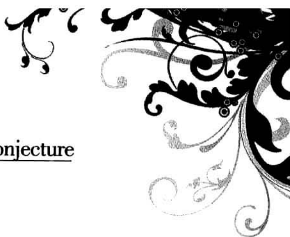
同样可定义向量空间的自由子集或约束子集的概念.

例子 在 $\mathbf{R}$ 上的函数空间中,单项式序列 $1,x,x^2,\cdots,x^n,\cdots$ 是自由的;序列 $\sin x,\cdots,\sin nx,\cdots$ 也是如此.

基底的定义 向量空间 E 中生成 E 的任何自由子集称为 E 的基底.

若 $B=(\boldsymbol{a}_i)_{i\in I}$ 是 E 的一个基底,每个 $\boldsymbol{x}\in E$ 唯一地表成 $\boldsymbol{x}=\sum\limits_{i\in I}x_i\boldsymbol{a}_i$,这里诸纯量 x_i 中,除有限个外均为零,x_i 称为 $\boldsymbol{x}$ 的第 i 个坐标,$x_i\boldsymbol{a}_i$ 称为 $\boldsymbol{x}$ 的第 i 分





量.

例如,在 $\mathbf{R}^n$ 中,n 个元素 $\boldsymbol{a}_i=(x_1,\cdots,x_n)(x_i=1$,其余为0) 的集合就是一个基底.在多项式空间中,单项式 $1,x,x^2,\cdots,x^n,\cdots$ 的集合是一个基底.

可以证明,每一向量空间至少有一个基底;下面仅就有限维空间证明这个事实.

有限维空间的基底

空间 E 如果包含一个生成它的有限子集,就叫有限维的这个有限子集也叫做有限生成组.

有限维空间的研究取决于下列定理:

定理1 若 G 是向量空间 E 的有限生成组,L 是 G 的自由子集,则存在 E 的一个基底 B,使得 $L\subset B\subset G$.

命 $L=L_0$,假设已确定了 G 的自由子集 L_p.若 L_p 不生成 E,则存在 G 的元素 $\boldsymbol{x}_{p+1}$ 不是 L_p 的元素的线性组合.G 的子集 $L_p\cup\{\boldsymbol{x}_{p+1}\}$ 显然是自由的,记为 L_{p+1}.

于是我们有 $L_0\subset L_1\cdots\subset L_p\subset L_{p+1}$ 这序列严格递增;因 $L_p\subset G$,而 G 有限,所以这序列有限.因此,存在一个 L_n 无后继的 L_{n+1};换言之,L_n 生成 E,这 L_n 就是所求的基底 B.

定理2 若 G 是 E 的有限生成组,L 是 E 的自由子集(不必含于 G),则存在 E 的一个基底 B,包含 L 并含于 $G\cup L$.

将定理1用于自由子集 L 与生成组($G\cup L$)即得.特别,E 的每个有限生成组均包含一个基底.

因此,每个有限维向量空间都具有有限基底.另外,这样的空间不能有无限基底,这一点并不明显,不过确实无误,更确切地说,利用定理2,使用归纳法可以证明,在有限维空间中,每个基底都是有限的,并且含有同样多个元素.

这个共同的个数称为空间的维数.注意,这个结论对于任意体 K 上的向量空间 E 也成立.不过,要紧的是要确切说明有关的体 K.事实上,若 E 是体 $\mathbf{C}$ 上的 n 维向量空间,空间 E 也可视为体 $\mathbf{R}$ 上的向量空间,但后者的维数不再是 n,而是 $2n$;例如,$\mathbf{C}$ 视为 $\mathbf{C}$ 自身上的向量空间是1维,但视为 $\mathbf{R}$ 上的向量空间便是2维.

上述结果有许多明显的推论,今列其最重要者如下:

1) 要使 $\mathbf{R}$ 上两个有限维向量空间 E 与 E' 同构,充要条件是它们的维数相等.特别,每个 n 维向量空间都与 $\mathbf{R}^n$ 同构.

2) 设 E 是 n 维向量空间,则:

a) E 的每个生成组至少含有 n 个元素;若恰好含有 n 个元,它就是一个基底.

此可以定义 E 与 $\mathbf{R}^n$ 之间的一一对应

$$p \leftrightarrow (a_1, a_2, \cdots, a_n)$$

这对应还保持加法与纯量乘法.所以,可以认为,E 与 $\mathbf{R}^n$ 是同一个代数结构的两个不同实现;这一事实,可以说成两个空间 E 与 $\mathbf{R}^n$ 同构.更一般有如下定义:

定义 设 E 与 E' 是同一个体 K 上的向量空间.每一个把 E 映成 E' 的一一映射 f 称为把 E 映成 E' 的同构映射,如果对任意的 $\boldsymbol{x}, \boldsymbol{y} \in E, \lambda \in K$ 有

$$f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}), f(\lambda \boldsymbol{x}) = \lambda f(\boldsymbol{x})$$

显然,把 E' 映成 E 的逆映射 f^{-1} 也有同样的性质.

如果在 E 与 E' 之间存在这样的同构映射,就说 E 与 E' 同构.

注意,从条件 $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$ 可推得对每个有理数 λ, $f(\lambda \boldsymbol{x}) = \lambda f(\boldsymbol{x})$;但对任意的纯量 x,不能推出这关系.因之,上述定义中的第二个条件才是主要的.

同构映射的例子

1) 设 E 是向量空间,λ 是纯量 $\neq 0$.把 E 映成 E 本身的映射 $\boldsymbol{x} \rightarrow \lambda \boldsymbol{x}$ 是同构映射,称为以 0 为中心、以 λ 为比例常数的位似.

2) 设 E_1 是 $[0,1]$ 上的连续函数的向量空间,E_2 是 $[0,1]$ 上当 $x = 0$ 时为零的连续可微函数的向量空间.由初等分析的一个定理知这两个空间同构.

3) 定义在 $\mathbf{R}^2$ 上的线性函数具有 $a_1x_1 + a_2x_2$ 的形式,因此,立即得:这类函数的集合是同构于 $\mathbf{R}^2$ 的向量空间.

4) $[0,1]$ 上的连续函数空间同构于 $(-\infty, +\infty)$ 上连续,并且当 $x \rightarrow +\infty$ 或 $x \rightarrow -\infty$ 时,$f(x)$ 趋于有限极限的函数 f 的空间.

5) 在 $\mathbf{R}$ 上连续且周期为 1 的周期函数的空间同构于定义在一圆周上的连续函数的空间.

向量空间的积

从 $\mathbf{R}$ 出发定义 $\mathbf{R}^n$ 所用的方法可以推广,这使我们可以从已知的空间构造出新的空间.

例如,设 E_1 与 E_2 是 $\mathbf{R}$ 上的两个向量空间,用 $E_1 \times E_2$ 表示偶对 $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)$ 的集合,这里 $\boldsymbol{x}_1 \in E_1, \boldsymbol{x}_2 \in E_2$.若 $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2), \boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2)$,我们令

$$\boldsymbol{x} + \boldsymbol{y} = (\boldsymbol{x}_1 + \boldsymbol{y}_1, \boldsymbol{x}_2 + \boldsymbol{y}_2), \lambda \boldsymbol{x} = (\lambda \boldsymbol{x}_1, \lambda \boldsymbol{x}_2)$$

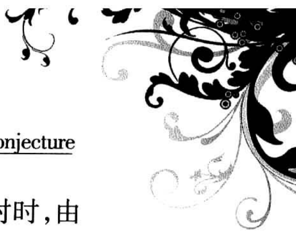
具有如上运算的集合 $E_1 \times E_2$ 就成为 $\mathbf{R}$ 上的向量空间,称为 E_1 与 E_2 的积.

同样可定义任意一族有限个乃至无限个向量空间的积.

例子 1) 定义在 $\mathbf{R}^2$ 上的形如 $h(x_1, x_2) = f(x_1) + g(x_2)$,且 $f(0) = g(0) = 0$ 的函数 $h(x_1, x_2)$ 的空间同构于定义在 $\mathbf{R}$ 上且 $f(0) = 0$ 的函数 f 所成空间的自乘.

2) $[0,1]$ 上连续可微函数的空间同构于 $\mathbf{R}$ 与 $[0,1]$ 上连续函数空间的积.





1)(传递性)若 f 把 E 映入 F,g 把 F 映入 G,则当 f 与 g 是线性映射时,由 $h(x)=g(f(x))$ 定义的映射 $h=g\circ f$ 是线性的.

2)若 f 是把 E 映入 F 的线性映射,则 E 的象 $f(E)$ 是 F 的向量子空间,F 的任一向量子空间 V 的逆象 $f^{-1}(V)$ 是 E 的向量子空间.

线性映射的秩

f 是把 E 映入 F 的线性映射,E 的象 $f(E)$ 的维数称为 f 的秩.下述命题常用来确定秩:

f 的秩等于 E 中与 $f^{-1}(0)$ 互补的子空间的维数.

事实上,令 G 是与 $f^{-1}(0)$(即是 $f(x)=0$ 的根的集合)互补的子空间.

对于每个 $x\in E$,有唯一分解 $\boldsymbol{x}=\boldsymbol{x}_0+\boldsymbol{x}_1$,这里 $\boldsymbol{x}_0\in f^{-1}(0)$,$\boldsymbol{x}_1\in G$.

我们有 $f(\boldsymbol{x})=f(\boldsymbol{x}_0)+f(\boldsymbol{x}_1)=f(\boldsymbol{x}_1)$,故 $f(G)=f(E)$.

然而,对每一 $\boldsymbol{x}\in G$,只要 $\boldsymbol{x}\neq 0$ 必有 $f(\boldsymbol{x})\neq 0$.故 f 在 G 上的限制是把 G 映成 $f(E)$ 的一一映射,所以这个限制是 G 与 $f(E)$ 之间的同构映射,由此命题得证.

推论 线性映射决不增大维数:当 E 有限维时,欲使 $f(E)$ 与 E 维数相等,充要条件是 f 是一一的.

这个推论刻画了线性映射的一个重要性质,这个性质并非每个连续映射都具有的,因为某些连续映射要增大维数.

生成组的象

设 f 是把 E 映入 F 的线性映射,X 是 E 的子集.

显然,X 生成的空间 $A(X)$ 的象 $f(A(X))$ 由 $f(x)$ 生成;当 X 生成 E 时,f 完全由它在 X 上的限制所确定.特别若 X 是 E 的基底,只要对这个基底的每个元素 $\boldsymbol{x}$ 确定了 $f(\boldsymbol{x})$.那么 f 就完全确定了.

4 线性形式,向量空间的对偶空间

设 E 是体 K 上的向量空间,把 E 映入 K 的任何线性映射,称为 E 上的一个线性形式.

E 上所有线性形式的集合,可以自然地具有(K 上)向量空间的结构,定义如下:

设 f 与 g 是两个线性形式,定义形式 $h=f+g$ 与 $k=\lambda f(\lambda\in K)$ 如下:对每个 $\boldsymbol{x}\in E$

$$h(\boldsymbol{x})=f(\boldsymbol{x})+g(\boldsymbol{x}),k(\boldsymbol{x})=\lambda f(\boldsymbol{x})$$

如上得到的向量空间称为 E 的对偶空间,记成 E^*.

对每个 $x'\in E^*$,$\boldsymbol{x}'(\boldsymbol{x})$ 是一个纯量,一般记为 $\langle\boldsymbol{x},\boldsymbol{x}'\rangle$.

对于每个固定的 $\boldsymbol{x}'$，$\langle \boldsymbol{x}, \boldsymbol{x}' \rangle$ 是 E 上的线性形式. 对于每个固定的 $\boldsymbol{x}$，$\langle \boldsymbol{x}, \boldsymbol{x}' \rangle$ 是 E^* 上的线性形式. 也说$\langle \boldsymbol{x}, \boldsymbol{x}' \rangle$是定义在 $E \times E^*$ 上的典范双线性形式.

坐标形式 · 对偶基

我们限于讨论有限维空间.

设 E 是 n 维向量空间，$(\boldsymbol{a}_i)_{i\in I}$ 是 E 的基底. E 的每个 $\boldsymbol{x}$ 唯一地写成 $\boldsymbol{x} = \sum \xi_i \boldsymbol{a}_i$.

把 E 映入 K 的映射 $\boldsymbol{x} \to \xi_i$ 是线性的，所以是一个线性形式，记作 $\boldsymbol{a}'_i$，称为关于所选取基底的第 i 个坐标形式.

$(\boldsymbol{a}'_i)_{i\in I}$ 构成E^* 的一个基底，称为 E 的基底$(\boldsymbol{a}_i)_{i\in I}$的对偶基. 事实上，对于每个 $\boldsymbol{x}' \in E^*$，我们有

$$\boldsymbol{x}'(\boldsymbol{x}) = \sum \boldsymbol{x}'(\xi_i \boldsymbol{a}_i) = \sum \xi_i \boldsymbol{x}'(\boldsymbol{a}_i) = \sum \boldsymbol{a}'_i(\boldsymbol{x})\boldsymbol{x}'(\boldsymbol{a}_i) \tag{1}$$

或者

$$\boldsymbol{x}' = \sum \boldsymbol{x}'(\boldsymbol{a}_i)\boldsymbol{a}'_i$$

这表明诸 $\boldsymbol{a}_i$ 生成 E^*；另一方面，诸 $\boldsymbol{a}_i$ 又是无关的，这是因为，关系 $\sum \lambda_i \boldsymbol{a}'_i = 0$ 等价于对每个族$(\xi_i)_{i\in I}$有$\sum \lambda_i \xi_i = 0$，所以，对每个 i 有$\lambda_i = 0$.

若令 $x'(\boldsymbol{a}_i) = \xi'_i$，即 $\boldsymbol{x}' = \sum \xi'_i \boldsymbol{a}'_i$，则公式(1) 表明

$$\langle \boldsymbol{x}, \boldsymbol{x}' \rangle = \sum \xi_i \xi'_i \tag{2}$$

当我们在 E 与E^* 中考虑两个对偶基时，公式(2) 就是$\langle \boldsymbol{x}, \boldsymbol{x}' \rangle$的标准形式. 这不外是欧氏空间上线性函数的熟知形式，例如，$f = ax + by + cz$. 因此有如下结论：

若 E 是 n 维空间，则 E^* 也是. 又若选取 E 与 E^* 的两个对偶基，则 $\langle \boldsymbol{x}, \boldsymbol{x}' \rangle = \sum_{i=1}^{n} \xi_i \xi'_i$.

5 线性方程

我们所处理的最一般的形式，可以使代数与分析中的许多问题变得形式简单.

定义 设 E 与 F 是同一体 K 上的两个向量空间. 每个形如 $f(\boldsymbol{x}) = \boldsymbol{y}_0$ 的方程称为线性方程，这里，f 是把 E 映入 F 的线性映射，$\boldsymbol{y}_0$ 是 F 中的已知元.

E 中每个使得$f(\boldsymbol{x}) = \boldsymbol{y}_0$ 的元素 $\boldsymbol{x}$ 叫这个方程的解.

方程 $f(\boldsymbol{x}) = 0$ 叫对应于已知方程 $f(\boldsymbol{x}) = \boldsymbol{y}_0$ 的齐次方程. $\boldsymbol{y}_0$ 称为已知方程的右端.

每个右端为零的线性方程称为齐次线性方程.





例子 1) E是定义在开圆盘$u^2+v^2<1$中,且具有各阶偏导数的函数构成的空间,F就是E.映射f是一个微分算子;例如,f是拉普拉斯算子$\frac{\partial^2}{\partial u^2}+\frac{\partial^2}{\partial v^2}$.这时,方程$f(\boldsymbol{x})=\boldsymbol{y}_0$不外就是非齐次的线性偏微分方程.

2) E是在闭圆盘$u^2+v^2\leqslant 1$上连续且在圆盘内部调和的函数构成的空间.F是相应圆周上的连续函数构成的空间.取映射f:对每个$\varphi\in E$,使φ在圆周上的限制与之对应.求函数φ,使得$f(\varphi)=y_0$,这里$y_0\in F$,也就是对于函数y_0求解狄利克雷问题.

3) E是$[0,1]$上的连续函数$x(t)$的空间,$F=E$.若$\phi(u,v)$表示正方形$0\leqslant u,v\leqslant 1$上的一连续函数,$f$是把$E$映入$F$的映射,定义为

$$f(x)(t)=\int_0^1\phi(t,\tau)x(\tau)\mathrm{d}\tau$$

方程$f(x)=b,b\in F$,不过是积分方程

$$\int_0^1\phi(\tau,\tau)x(\tau)\mathrm{d}\tau=b(t)$$

线性方程的解的形式

定理 设$f(\boldsymbol{x})=\boldsymbol{y}_0$是一线性方程.若$\boldsymbol{x}_0$是这方程的任一解,那么,把对应齐次方程的任意解与$\boldsymbol{x}_0$相加就得到$f(\boldsymbol{x})=\boldsymbol{y}_0$的全部解.因此,解的集合是仿射簇$\boldsymbol{x}_0+f^{-1}(\boldsymbol{0})$,由$E$的向量子空间$f^{-1}(\boldsymbol{0})$平移$\boldsymbol{x}_0$而得.

欲使方程至多只有一个解,充要条件是$f^{-1}(\boldsymbol{0})=\{\boldsymbol{0}\}$,所以,$f$是把$E$映成$f(E)$的同构映射.

上述结论的证明是明显的.

方程组 一个线性方程常可表成线性方程组形式.

设E是向量空间,$(F_i)_{i\in I}$是一族有限或无限多个向量空间,对每一$i\in I$,令f_i是把E映入F_i的线性映射.

方程族$(f_i(\boldsymbol{x})=\boldsymbol{b}_i)_{i\in I},\boldsymbol{b}_i\in F_i$,称为线性方程组,这个方程组的解是指该组内所有方程的公共解.

这样的方程组等价于一个线性方程.事实上,设F是诸F_i的积空间;$\boldsymbol{b}=(\boldsymbol{b}_i)$,$f$是把$E$映入$F$的映射,它的分量是$f_i$.方程$f(\boldsymbol{x})=\boldsymbol{b}$就等价于所给的方程组.

反之,一个线性方程也可以换成方程组.例如,设F是q维空间,$(\boldsymbol{b}_j)_{j\in J}$是$F$的基底;对每个$\boldsymbol{y}\in F$,有$\boldsymbol{y}=\sum\beta_j\boldsymbol{b}_j$.

于是,给出把空间E映入F的一个线性映射,就等于给出q个线性形式f_j,使得对每个$\boldsymbol{x}\in E$有

$$f(\boldsymbol{x})=\sum f_j(\boldsymbol{x})\boldsymbol{b}_j$$

因此,每个方程 $f(\boldsymbol{x}) = \boldsymbol{y}$ 等价于一组 q 个方程 $f_j(\boldsymbol{x}) = \beta_j$,后者还可写成 $\langle \boldsymbol{x}, f_j \rangle = \beta_j$.这样的方程组称为纯量方程组.

现在我们再假定 E 也是有限维的.

在有限维空间上的纯量方程组

定理 设 E 是 n 维向量空间;$(\boldsymbol{x}'_i)_{i\in I}$ 是 E^* 的 r 个元素(即 E 上的 r 个线性形式)的自由族,则对任意的纯量 η_i,纯量组

$$\langle \boldsymbol{x}, \boldsymbol{x}'_i \rangle = \eta_i \quad (i \in I)$$

有解.解的集合形如 $\boldsymbol{x}_0 + V$,这里 V 是 E 的 $n - r$ 维向量子空间.

事实上,$\boldsymbol{x}'_i$ 属于 E^* 的一个基底 B^*,这个基底是 E 的一个基底 B 的对偶基底.

对于 B^*,每个 $\boldsymbol{x}'_i$ 所有的坐标,除一个外都为零,所以在对偶基 B 与 B^* 下,$\langle \boldsymbol{x}, \boldsymbol{x}'_i \rangle = \xi_i$,因此上面的纯量方程组写成

$$\xi_i = \eta_i \quad (i \in I)$$

于是这个方程组对于 $i \in I$ 有解 $\xi_i = \eta_i$.对于 $i \notin I$ 以任意 ξ_i 为解.因之,定理中的 V 是 E 中使得坐标 $(\xi_i)_{i\in I}$ 全为零的向量组成的子空间,其维数正好是 $n - r$.

为了研究任意的纯量方程组,引入下面定义:

E^* 中由族 $(\boldsymbol{x}'_i)_{i\in I}$ 所生成的子空间的维数(即线性无关的 $\boldsymbol{x}'_i$ 的最大数目),称为纯量方程组 $\langle \boldsymbol{x}, \boldsymbol{x}'_i \rangle = \eta_i (i \in I)$ 的秩.

利用前面的定理,可以证明:

定理 设 $\langle \boldsymbol{x}, \boldsymbol{x}_i \rangle = \eta_i (i \in I)$ 是 n 维向量空间 E 上秩为 r 的纯量方程组.要使它至少有一解,充要条件是对每个零线性组合 $\sum \lambda_i \boldsymbol{a}'_i = 0$,也有 $\sum \lambda_i \eta_i = 0$.

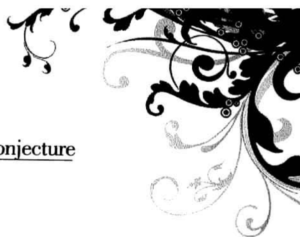
当这个方程组有解 $\boldsymbol{x}_0$ 时,解的集合形如 $\boldsymbol{x}_0 + V$,这里 V 是 E 的 $n - r$ 维向量子空间.

推论 1)要使 n 维向量空间 E 上的齐次纯量方程组有非零解,充要条件是它的秩 $r < n$.

2)要使 n 维空间上的一组 n 个纯量方程组有一且仅有一解(克莱姆方程组),充要条件是对应的齐次组只有零解.

为了进一步研究纯量方程组,特别是为了给出可以确定它的秩的方便方法,要用到行列式的理论,本节就从略了.





第四讲　线性映射与矩阵[①]

这期讲座的前面几讲中,研究了主要的代数结构:群的结构,环的结构,体的结构,以及向量空间的结构.我这一讲是讨论把一个向量空间映入另一个向量空间的线性映射及其矩阵表示,这一讲里我们将碰到上述各种概念的一些例子.

本讲的主题正好相当于前不久所谓的线性变换的理论;老实说,现代观点并未产生任何实质上新颖的结果.我要讲的东西,你们大家都知道.现代观点以及由此产生的数学活动规律的变化,其所以有价值,正是由于方法的谐调一致,即在于概括的威力.

这时的线性变换理论中表达方式和标号细腻入微,在其浓云迷雾的掩盖下,基本结果和证明的主导思想便很难突出,行列式以及由此建立起来的线性方程组理论似乎起着主要的作用.

按照我们新的搞法,喜欢尽可能少用表达式,而对向量和映射加以直接的、内在的论证,清楚地突出奠定理论基础的那些重要的代数结构,而且不再使用行列式;我们觉得行列式是一种工具,用得过分,就会掩盖那些基本代数事实简单明白的面貌.在我的讲演里,几乎不用行列式,只是讲特征值的那部分是个例外;那里不用行列式也是行之有效的,不过这一点我就不多啰唆了.我仍然假定大家知道古典的线性方程理论,竭力选择各种不同的证明,以显示近世代数证明方式的运用灵活与丰富多彩.

我要给你们讲的理论,无疑会使你们感到别扭而不会让你们为之倾倒,因为你们也和我一样,本来都习惯于另外一套东西.请你们不要怪我过分"抽象":这个颇为悦耳的不逊之词究竟是什么意思,是很难说清楚的;而散布在黑板上的粉笔灰的数量也不能说明我讲的内容的具体程度.具体的东西往往正好是我们习以为常的东西,即是在我们个人的成长过程中占主导地位的东西.对于我这一代很多数学家来说,近世代数在我们掌握之前也显得抽象.而在我们更年轻的同事中那些(恕我冒说一句)从小受到向量空间熏陶的人看来,近世代数正是光彩夺目的具体之物.

复习　集合 E 上关于一个可换体 K(纯量体)的向量空间结构由下列两个

① 原作者 A. Lichnerowicz(法兰西学院教授).

合成法则确定:

1° E 的一个内合成法则,用加法表示,在 E 上定义了一个阿贝尔群的结构;

2° 一个称为“用纯量乘”的合成法则:对于 $\alpha \in K, \boldsymbol{x} \in E$ 相应地给出 $\alpha\boldsymbol{x} \in E$,具有如下性质:对任意的 $\alpha, \beta \in K, \boldsymbol{x}, \boldsymbol{y} \in E$ 有

$$1 \cdot \boldsymbol{x} = \boldsymbol{x}, \alpha(\beta\boldsymbol{x}) = (\alpha\beta)\boldsymbol{x}$$

$$(\alpha + \beta)\boldsymbol{x} = \alpha\boldsymbol{x} + \beta\boldsymbol{x}, \alpha(\boldsymbol{x} + \boldsymbol{y}) = \alpha\boldsymbol{x} + \alpha\boldsymbol{y}$$

这里所考虑的向量空间都假定为有限维的,维数是指 E 中自由向量的最大个数(或者线性无关向量的最大个数).

E 的向量子空间是 E 的子集 V,使得对任意的 $\boldsymbol{x}, \boldsymbol{y} \in V, \alpha \in K$,有 $\boldsymbol{x} + \boldsymbol{y}$, $\alpha\boldsymbol{x} \in V$. E 的一组向量的线性组合是 E 的一个子空间的元素,这个子空间称为“由该向量组生成”.向量组的秩是其中所含自由向量的最大个数,也就是这组向量生成的子空间的维数.

给了 E 的子空间 V,则存在子空间 U,使得 E 的每个向量都能唯一地表成 U 的一个向量与 V 的一个向量的和.特别有 $U \cap V = 0$; U 叫做 V 在 E 中的补子空间.

1 线性映射概要

线性映射的概念

a) 设有两个向量空间 E 与 F,维数分别为 n 与 p,确定在同一纯量体 K 上(例如实数体或复数体).

定义 把 E 映入 F 的一个映射 f 叫做线性映射,如果满足下列两个条件:

1) 对任意 $\boldsymbol{x}, \boldsymbol{y} \in E$ 有 $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

2) 对任意 $\boldsymbol{x} \in E, \alpha \in K$ 有 $f(\alpha\boldsymbol{x}) = \alpha f(\boldsymbol{x})$.

把 E 映成 F 的线性映射 f 叫做同态映射;若 f 还是一一的,则 f 叫把 E 映成 F 的同构映射.

E 上的线性形式,就是一个线性映射,把 E 映入 K 本身所构成的一维向量空间.若该形式不恒为零,则它还是把 E 映成 K 的同态映射.初等几何中讲到很多线性映射:在初等几何的空间中给定坐标系 $Oxyz$,取所有(自由)向量的空间为 E 平面 xOy 上向量构成的空间为 F,平行于 Oz 的射影为映射 f,这是把 E 映成 F 的映射.

同样,仍取 E 如上, $F = E$.中心为 O,比率为 λ 的位似定义一个把 E 映成自身的一一映射.

b) E 的向量由 f 产生的象是 F 中的向量,它们构成 F 的子集 $f(E)$ 是 F 的一个向量子空间.事实上,若 $(\boldsymbol{l}_i)$ 是 E 的任一基底,因 f 线性,故相应于向量





$$\boldsymbol{x} = \sum_{i=1}^{n} x^i \boldsymbol{l}_i \quad (x^i \in K) \tag{1}$$

得到 F 的向量

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} x^i f(\boldsymbol{l}_i) \tag{2}$$

因此,$f(E)$ 的每个元素均可表成 $f(\boldsymbol{l}_i)$ 的线性组合.反之,F 的每个向量,如果可以表为系数是 (x^i) 的这种线性组合,就是(1)由 f 产生的象,从而 $f(E)$ 是 F 中由 $f(\boldsymbol{l}_i)$ 生成的子空间.

$f(E)$ 的维数叫做线性映射 f 的秩;这个秩 r 正好就是 F 的那组 n 个向量 $f(\boldsymbol{l}_i)$ 的秩,故 $r \leqslant n,p$.

c) 若把 E 映成 $f(E)$ 的映射是一一的,则 $f(\boldsymbol{x}) = 0$ 蕴涵 $\boldsymbol{x} = \boldsymbol{0}$.若式(2)确定的 $f(\boldsymbol{x})$ 为零,则式(1)确定的 $\boldsymbol{x}$ 为零,所以 $x^i = 0$.因此,诸 $f(\boldsymbol{l}_i)$ 构成自由组,$r = n$,或

$$\dim f(E) = \dim E$$

反之,考虑两个同维数的空间 E 与 F.容易构造它们之间的一个同构映射:若 $(\boldsymbol{l}_i)$ 是 E 的基底,$(\boldsymbol{f}_i)$ 是 F 的基底,设 f 把 $\boldsymbol{x} = \sum x^i \boldsymbol{l}_i \in E$ 映成 $f(\boldsymbol{x}) = \sum x^i \boldsymbol{f}_i \in F$.立即可以验明,这样就定义了一个把 E 映成 F 的一一线性映射 f,即同构映射.因此,两个向量空间同构的充要条件是它们维数相同.

向量子空间的象

a) 线性映射 f 有个基本性质,即保持向量空间的结构.确切地讲,若 V 是 E 的向量子空间,则 V 的向量由 f 产生的象是 F 的向量,它们构成的 F 的一个子集 $f(V)$ 成为 F 的向量子空间.只需将 f 限制在 V 上,应用刚才对 $f(E)$ 得到的结果即可.若 f 一一,则 $\dim V = \dim f(V)$;向量组的秩在同构映射下不变.

若 W 是 $f(E)$ 的子空间,$f^{-1}(W)$ 是 E 的向量集合,这些向量由 f 产生的象属于 W.显然,$f^{-1}(W)$ 仍是 E 的子空间;事实上,若 $\boldsymbol{x},\boldsymbol{y} \in f^{-1}(W)$,则 $f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}) \in W$,故 $\boldsymbol{x} + \boldsymbol{y} \in f^{-1}(W)$;同样,若 $\alpha \in K$,则 $f(\alpha\boldsymbol{x}) = \alpha f(\boldsymbol{x}) \in W$,故 $\alpha\boldsymbol{x} \in f^{-1}(W)$.

特别,若取 $W = 0$,则 $f^{-1}(0)$ 是 E 的子空间.

b) 关于这个子空间,我们将建立如下定理:

定理 向量空间 $f(E)$ 同构于 $f^{-1}(0)$ 在 E 中的任何补子空间.

设 L 是 $f^{-1}(0)$ 在 E 中的补子空间.对于每个元素 $z \in L$,相应地得到 $f(E)$ 的元素 $f(z)$.易见 $f(E) = f(L)$;事实上,若 $f(\boldsymbol{x})(\boldsymbol{x} \in E)$ 是 $f(E)$ 的任意元素,考虑分解

$$\boldsymbol{x} = \boldsymbol{y} + z \quad (\boldsymbol{y} \in f^{-1}(0), z \in L)$$

我们有 $f(\boldsymbol{x}) = f(\boldsymbol{y}) + f(z) = f(z)$(因为 $f(\boldsymbol{y}) = 0$).

如上定义的映射是把 L 映成$f(E)$ 的一一映射，因为，若 $f(z)=0$，则 $z\in L\cap f^{-1}(0)$ 是零元素. 这样，我们就定义了把 L 映成$f(E)$ 的一一线性映射，故 L 与 $f(E)$ 同构. 由此推出

$$\dim f^{-1}(0)=n-r$$

特别，要使 f 是把E 映成$f(E)$ 的同构映射，充要条件是 $r=n$. 要使 f 把E 映成 F，充要条件是 $r=p$. 要使 f 把E 映成F 的同构，充要条件是 $r=n=p$.

把 E 映入 F 的映射构成的向量空间

对于把 E 映入F 的线性映射构成的集合$L(E,F)$，我们可以把 E 上的线性形式的集合中引入的合成法则加以推广，在 $L(E,F)$ 上引入两个合成法则如下：

1° 对于已知的$f,g\in L(E,F)$ 我们可以相应地得到一个映射，记为$f+g$，定义如下：对任意的 $\boldsymbol{x}\in E$

$$f+g:\boldsymbol{x}\rightarrow f(\boldsymbol{x})+g(\boldsymbol{x})$$

这映射显然是线性的.

2° 对于已知的$\alpha\in K,f\in L(E,F)$. 我们可以相应地得到一个映射，记为 αf，定义如下：对任意的 $\boldsymbol{x}\in E$

$$\alpha f:\boldsymbol{x}\rightarrow f(\alpha\boldsymbol{x})=\alpha f(\boldsymbol{x})$$

这 αf 显然也是线性的.

容易验明，这两个法则总起来在 $L(E,F)$ 中确定了向量空间的结构. 于是我们有了线性映射的相加，以及线性映射与纯量的乘法. 零映射使得与所有 $\boldsymbol{x}\in E$ 对应的元素都是 $0\in F$；f 的反映射使得对于每个 $\boldsymbol{x}\in E$ 相应地给出 $f(\boldsymbol{x})$ 的反向量.

两个或多个线性映射的积

考虑三个向量空间 E,F,G. 维数分别是 n,p,q. 设 f 是把E 映入F 的线性映射，g 是把F 映入G 的线性映射. 今考虑把 E 映入G 的一个映射，定义如下：若 $\boldsymbol{x}\in E$，则按 f 得到F 的元素 $\boldsymbol{y}=f(\boldsymbol{x})$，然后按 g 得到 G 的元素 $\boldsymbol{z}=g(\boldsymbol{y})=g(f(\boldsymbol{x}))$. 这个把 E 映入 G 的映射记为 $g\circ f$. 它是线性的，因为对于 $\alpha\in K,\boldsymbol{x},\boldsymbol{x}_1,\boldsymbol{x}_2\in E$ 有

$$g(f(\boldsymbol{x}_1+\boldsymbol{x}_2))=g(f(\boldsymbol{x}_1)+f(\boldsymbol{x}_2))=g(f(\boldsymbol{x}_1))+g(f(\boldsymbol{x}_2))$$

$$g(f(\alpha\boldsymbol{x}))=g(\alpha f(\boldsymbol{x}))=\alpha g(f(\boldsymbol{x}))$$

线性映射 $g\circ f$ 叫做线性映射f 与g 的合成或积. 这个乘显然可以推广到两个以上的线性映射，而且按定义显然是结合的，关于($L(E,F)$ 或 $L(F,G)$) 线性映射的加法是分配的.





2 矩阵表示

线性映射的矩阵表示

a) 取 n 维向量空间 E,p 维向量空间 F,在 E 中选一基底($\boldsymbol{l}_i$).这一基底取定之后,对于每个线性映射 $f \in L(E,F)$,相应地得到 F 的一组 n 个向量($f(\boldsymbol{l}_i)$);反之,对 F 的任何一组 n 个向量($\boldsymbol{f}_i$),也相应地有一个把 E 映入 F 的映射 f,即是把

$$\boldsymbol{x} = \sum_{i=1}^{n} x^i \boldsymbol{l}_i \in E$$

映成

$$\boldsymbol{y} = \sum_{i=1}^{n} x^i \boldsymbol{f}_i \in F$$

的映射.这个映射是线性的,并且使得 $f(\boldsymbol{l}_i) = \boldsymbol{f}_i$.

因此,在 E 中选定一基底,我们能定义一个一一映射,把 $L(E,F)$(把 E 映入 F 的线性映射的集合)映成 F 的所有 n 元向量组构成的集合.

b) 为了确定这样一组中的所有向量,最方便的是按照 F 的一个基底引入它们的分量.设有 F 的基底($\boldsymbol{\varepsilon}_\mu$)($\mu = 1,2,\cdots,p$),将 $f(\boldsymbol{l}_i)$ 用这基底表出

$$\begin{cases} f(\boldsymbol{l}_1) = a_1^1\boldsymbol{\varepsilon}_1 + a_1^2\boldsymbol{\varepsilon}_2 + \cdots + a_1^p\boldsymbol{\varepsilon}_p \\ f(\boldsymbol{l}_2) = a_2^1\boldsymbol{\varepsilon}_1 + a_2^2\boldsymbol{\varepsilon}_2 + \cdots + a_2^p\boldsymbol{\varepsilon}_p \\ \vdots \\ f(\boldsymbol{l}_n) = a_n^1\boldsymbol{\varepsilon}_1 + a_n^2\boldsymbol{\varepsilon}_2 + \cdots + a_n^p\boldsymbol{\varepsilon}_p \end{cases} \tag{3}$$

这样我们一共引入了 K 的 np 个数:把它们排成一个矩形表

$$\boldsymbol{A} = \begin{pmatrix} a_1^1 & a_1^2 & \cdots & a_1^p \\ a_2^1 & a_2^2 & \cdots & a_2^p \\ \vdots & \vdots & & \vdots \\ a_n^1 & a_n^2 & \cdots & a_n^p \end{pmatrix} = (a_i^\mu)$$

$$(i = 1,2,\cdots,n;\mu = 1,2,\cdots,p)$$

这样一个表用字母 $\boldsymbol{A}$(或别的大写字母)表示,称为 n 行 p 列矩阵(或 $n \times p$ 矩阵).

因此,在 E 与 F 中分别选定了基底后,对于每个线性映射,通过(3)得到相应的 $n \times p$ 矩阵,称为 f 在所选定基底下的表示矩阵.反之,设有 K 的元素组成的 $n \times p$ 矩阵,在同一行上的元素,是某个向量 $\boldsymbol{f}_i$ 在 F 的已知基底下的分量,这样就得到 F 的 n 个向量,根据 a),这些向量定义一个线性映射 f.于是有

定理 设选定 E 与 F 的基底,公式(3)定义了一个一一映射,把 $L(E,F)$

映成所有 $n\times p$ 矩阵的集合.

c) 考虑向量空间 K^p,它的元素是 K 中 p 个数组成的序列,K^p 有一标准基底(基底$(1,0,\cdots,0),(0,1,0,\cdots,0),\cdots,(0,0,\cdots,1)$).设有一 $n\times p$ 矩阵 $\boldsymbol{A}$,$\boldsymbol{A}$ 的每一行确定 K^p 的一个向量,称为行向量.这一组 n 个行向量的秩称为矩阵 $\boldsymbol{A}$ 的秩.

F 的基底给定后,就确定 K^p 与 F 之间的一个同构映射,使得互相对应的向量有相同的分量.对于任意选取的 E 与 F 的基底.$\boldsymbol{A}$ 的秩等于诸行向量确定的 F 中那一组 n 个向量的秩,所以 $\boldsymbol{A}$ 的秩就是以 $\boldsymbol{A}$ 为表示矩阵的映射 f 的秩.因此,每个矩阵 $\boldsymbol{A}$ 与其所表示的线性映射 f 有相同的秩.

正如常用向量的分量去处理向量,或用点的坐标处理点一样,用线性映射的表示矩阵来处理映射也往往很方便,其结果就是地地道道的有关线性映射的"解析几何".

矩阵表示的例子

a) 投影矩阵

在通常的几何空间中,给了三个单位向量 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$,两两正交.命 E 为通常的几何向量组成的向量空间,F 是 $\boldsymbol{i},\boldsymbol{j}$ 生成的子空间,f 是 E 在 F 上的垂直投影,这是把 E 映成 F 的线性映射;E 有基底$(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$,$F$ 有基底$(\boldsymbol{i},\boldsymbol{j})$,我们有

$$\begin{cases}f(\boldsymbol{i})=\boldsymbol{i}\\f(\boldsymbol{j})=\boldsymbol{j}\\f(\boldsymbol{k})=\boldsymbol{0}\end{cases}$$

由此表示矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$$

它的秩是 2,$f(E)$ 的维数 $=$ F 的维数.

b) 位似矩阵

E 同上,F 与 E 同,f 是把 E 映成自身的一个线性映,由比率为 λ 的一个位似确定;E 与 F 有基底$(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$

$$\begin{cases}f(\boldsymbol{i})=\lambda\boldsymbol{i}\\f(\boldsymbol{j})=\lambda\boldsymbol{j}\\f(\boldsymbol{k})=\lambda\boldsymbol{k}\end{cases}$$

由此表示矩阵

$$\boldsymbol{A}=\begin{pmatrix}\lambda&0&0\\0&\lambda&0\\0&0&\lambda\end{pmatrix}$$





它的秩是 3.

矩阵代数:加法以及与纯量的乘法

用 $M(n,p)$ 表示由 K 的元素构成的 $n \times p$ 矩阵的集合.我们要在 $M(n,p)$ 上定义一个关于 K 的向量空间结构.

a) $M(n,p)$ 中两个元素的加法.设 A 与 B 是两个由 K 的元素构成的 $n \times p$ 矩阵.若 E 与 F 是 K 上的两个向量空间,维数分别是 n 与 p,基底分别为$(\boldsymbol{l}_i)$ 与 $(\boldsymbol{\varepsilon}_\mu)$,$\boldsymbol{A}$,$\boldsymbol{B}$ 是把 E 映入 F 的两个线性映射 f 与 g 在此基底下的表示矩阵.

我们来讨论映射 $f+g$ 在如上基底下的表示矩阵 $\boldsymbol{S}$.若 $\boldsymbol{A}=(a_i^p)$,$\boldsymbol{B}=(b_i^p)$,我们有

$$f(\boldsymbol{l}_i)=\sum_\mu a_i^\mu \boldsymbol{\varepsilon}_\mu,\ g(\boldsymbol{l}_i)=\sum_\mu b_i^\mu \boldsymbol{\varepsilon}_\mu$$

所以
$$(f+g)(\boldsymbol{l}_i)=f(\boldsymbol{l}_i)+g(\boldsymbol{l}_i)=\sum_\mu (a_i^\mu+b_i^\mu)\boldsymbol{\varepsilon}_\mu$$

$\boldsymbol{S}$ 的元素为

$$s_i^\mu = a_i^\mu + b_i^\mu \tag{4}$$

$n \times p$ 矩阵 $\boldsymbol{S}$,其元素由(4) 给出,称为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和

$$\boldsymbol{S}=\boldsymbol{A}+\boldsymbol{B}$$

这样定义的合成法则使 $M(n,p)$ 具有阿贝尔群的结构.矩阵 $\boldsymbol{O}$ 是所有元素都为零的矩阵;把一个矩阵的元素都换成其反元就得到该矩阵的反矩阵.

b) 与纯量的积,记号不变,$\alpha \in K$,讨论映射 αf 的表示矩阵 $\boldsymbol{P}$.因

$$f(\boldsymbol{l}_i)=\sum_\mu a_i^\mu \boldsymbol{\varepsilon}_\mu,\ \alpha f(\boldsymbol{l}_i)=\sum_\mu \alpha a_i^\mu \boldsymbol{\varepsilon}_\mu$$

故 $\boldsymbol{P}$ 的元素为

$$p_i^\mu = \alpha a_i^\mu \tag{5}$$

容易验明,如上定义的两个法则使 $M(n,p)$ 具有向量空间的结构.

c) 在 E 与 F 中各选定一基底,考虑把 $L(E,F)$ 映成 $M(n,p)$ 的一个一一映射 ϕ:对任何线性映射 $f \in L(E,F)$ 使它的表示矩阵 $\boldsymbol{A}=\phi(f)$ 与之相应.根据上面加法以及与纯量的积的定义有

$$\phi(f+g)=\phi(f)+\phi(g),\ \phi(\alpha f)=\alpha\phi(f)$$

所以,ϕ 是把向量空间 $L(E,F)$ 映成向量空间 $M(n,p)$ 的同构映射.

容易构造 $M(n,p)$ 的一个基底:设 $\boldsymbol{E}_\mu^i$ 是 $n \times p$ 矩阵,它的每个元素除去 $e_i^\mu=1$ 外,其余为 0.显然,若 $\boldsymbol{A}=(a_i^\mu)$,则

$$\boldsymbol{A}=\sum_{i,\mu} a_i^\mu \boldsymbol{E}_\mu^i$$

这样,每个 $n \times p$ 矩阵都可以表成诸 $\boldsymbol{E}_\mu^i$ 的线性组合.而且对 $\boldsymbol{O}$ 矩阵,所有的系数都必须是 0.诸 $\boldsymbol{E}_\mu^i$ 构成 $M(n,p)$ 的一基底,它由 np 个元素(矩阵)组成.因

此,彼此同构的向量空间 $M(n,p)$ 与 $L(E,F)$ 的维数都是 np.

矩阵的积

从线性映射的合成或积出发,在某些情形可以定义矩阵的积.设 E,F,G 是 K 上的三个向量空间,维数分别为 n,p,q,分别选定基底$(\boldsymbol{l}_i)$,$(\boldsymbol{\varepsilon}_\mu)$,$(\boldsymbol{\eta}_\lambda)$.

对于所选的这些基底,命 $n\times p$ 矩阵$\boldsymbol{A}$ 是把E 映入F 的映射f 的表示矩阵,$p\times q$ 矩阵$\boldsymbol{B}$ 是把F 映入G 的映射g 的表示矩阵.现在讨论把 E 映入G 的映射 $g\circ f$ 的表示矩阵$\boldsymbol{D}$;$\boldsymbol{D}$ 是$n\times q$ 矩阵.若

$$f(\boldsymbol{l}_i)=\sum_\mu a_i^\mu\boldsymbol{\varepsilon}_\mu,g(\boldsymbol{\varepsilon}_\mu)=\sum_\lambda b_\mu^\lambda\boldsymbol{\eta}_\lambda$$

则

$$g(f(\boldsymbol{l}_i))=\sum_\mu a_i^\mu g(\boldsymbol{\varepsilon}_\mu)=\sum_{\mu,\lambda}a_i^\mu b_\mu^\lambda\boldsymbol{\eta}_\lambda$$

即

$$g(f(\boldsymbol{l}_i))=\sum_\lambda\left(\sum_\mu a_i^\mu b_\mu^\lambda\right)\boldsymbol{\eta}_\lambda$$

故矩阵 $\boldsymbol{D}$ 的元素为

$$d_i^\lambda=\sum_\mu a_i^\mu b_\mu^\lambda \tag{6}$$

矩阵 $\boldsymbol{D}$ 叫做$\boldsymbol{A}$ 乘以$\boldsymbol{B}$ 的积,记成 $\boldsymbol{AB}$.仅当 $\boldsymbol{A}$ 的列数等于$\boldsymbol{B}$ 的行数时,乘积 $\boldsymbol{AB}$ 才能进行,乘积 $\boldsymbol{AB}$ 是n($\boldsymbol{A}$ 的行数)$\times q$($\boldsymbol{B}$ 的列数)矩阵.下列性质是显然的.

1)矩阵的积是结合的:若 $\boldsymbol{A}$ 是$n\times p$ 矩阵,$\boldsymbol{B}$ 是$p\times q$ 矩阵,$\boldsymbol{C}$ 是$q\times r$ 矩阵,则

$$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$$

这显然是因为各个矩阵所表示的线性映射的积满足结合律.

2)矩阵的积关于加法是右分配与左分配的:若 $\boldsymbol{A},\boldsymbol{A}_1$ 与 $\boldsymbol{A}_2$ 是 $n\times p$ 矩阵,$\boldsymbol{B},\boldsymbol{B}_1$ 与 $\boldsymbol{B}_2$ 是 $p\times q$ 矩阵,则

$$(\boldsymbol{A}_1+\boldsymbol{A}_2)\boldsymbol{B}=\boldsymbol{A}_1\boldsymbol{B}+\boldsymbol{A}_2\boldsymbol{B},\boldsymbol{A}(\boldsymbol{B}_1+\boldsymbol{B}_2)=\boldsymbol{AB}_1+\boldsymbol{AB}_2$$

因为相应的线性映射有同样的关系.

3)设 $\boldsymbol{A}$ 是$n\times p$ 矩阵,$\boldsymbol{B}$ 是$p\times q$ 矩阵,我们可作积 $\boldsymbol{AB}$,但只有 $n=q$ 时才能作积 $\boldsymbol{BA}$,这时 $\boldsymbol{BA}$ 是$p\times q$ 矩阵,而 $\boldsymbol{AB}$ 是$n\times n$ 矩阵.即使还有 $n=p$,我们也容易举例说明 $\boldsymbol{AB}$ 与$\boldsymbol{BA}$ 一般是不同的,所以,乘积不是可换的.

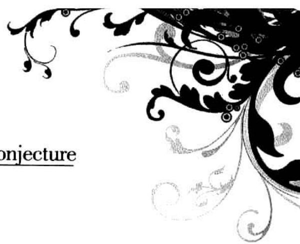
积的例子

a)考虑直角坐标平面 xOy,$\boldsymbol{i},\boldsymbol{j}$ 是坐标轴上的单位向量;E,F,G 都是$\boldsymbol{i},\boldsymbol{j}$ 所生成的向量空间,取f 为旋转α 角的映射,g 为关于Ox 轴的对称映射,则对基底$(\boldsymbol{i},\boldsymbol{j})$而言

$$\begin{cases}f(\boldsymbol{i})=\boldsymbol{i}\cos\alpha+\boldsymbol{j}\sin\alpha\\f(\boldsymbol{j})=-\boldsymbol{i}\sin\alpha+\boldsymbol{j}\cos\alpha\end{cases}$$

的表示矩阵为





$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}, \begin{cases} g(\boldsymbol{i}) = \boldsymbol{i} \\ g(\boldsymbol{j}) = -\boldsymbol{j} \end{cases}$$

的表示矩阵为

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

所以

$$\boldsymbol{BA} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

$$\boldsymbol{AB} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ -\sin\alpha & -\cos\alpha \end{pmatrix}$$

几何上看,$\boldsymbol{BA}$ 表示先关于 Ox 轴作对称然后旋转 α 角,即关于倾角为 $\frac{\alpha}{2}$ 的直线作对称;$\boldsymbol{AB}$ 表示关于倾角为 $-\frac{\alpha}{2}$ 的直线对称.

b) 再考虑矩阵表示这一小节中例 a) 的投影矩阵,但视为把 E 映入 E 的映射.它在$(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})$ 下的表示矩阵为

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

几何上看,$f \circ f = f$,由此推出 $\boldsymbol{AA} = \boldsymbol{A}$,这结果也可以直接验证.

一个向量关于基底的矩阵表示

设 E 是体 K 上的向量空间,把 K 视为一维向量空间,其标准基底是 $\boldsymbol{l}$ 对于把 K 映入 E 的每个线性映射 ξ,相应地有 E 的向量 $\boldsymbol{x} = \xi(\boldsymbol{l})$,反之,对于 E 的每个向量 $\boldsymbol{x}$,也相应地有一个把 K 映入 E 的映射 ξ

$$\xi: \alpha \rightarrow \alpha\boldsymbol{x} \quad (\alpha \in K)$$

因此我们在 E 与 $L(K,E)$ 之间定义了一个同构映射.

在 E 中选定基底$(\boldsymbol{l}_i)$,则

$$\xi(\boldsymbol{l}) = \boldsymbol{x} = x^1\boldsymbol{l}_1 + x^2\boldsymbol{l}_2 + \cdots + x^n\boldsymbol{l}_n$$

从而 ξ 关于所选基底的表示矩阵是 $1 \times n$ 矩阵

$$\boldsymbol{X} = (x^1 \quad x^2 \quad \cdots \quad x^n)$$

这个矩阵叫向量 $\boldsymbol{x}$ 关于基底$(\boldsymbol{l}_i)$ 的表示矩阵.

设 F 是 K 上的另一向量空间,f 是把 E 映入 F 的线性映射.映射 $f \circ \xi$ 是把 K 映入 F 的线性映射,相当于向量 $f(\boldsymbol{x})$:若 $\boldsymbol{A}$ 是 f 关于基底$(\boldsymbol{l}_i)$ 的表示矩阵,$(\boldsymbol{\varepsilon}_\mu)$ 是 F 的基底,则 $f \circ \xi$ 关于 $\boldsymbol{l}$ 与$(\boldsymbol{\varepsilon}_\mu)$ 的表示矩阵是 $\boldsymbol{XA}$,这个 $1 \times p$ 矩阵是 F 的向量 $f(\boldsymbol{x})$ 关于基底$(\boldsymbol{\varepsilon}_\mu)$ 的表示矩阵.

这样,一个向量空间的向量可以表为一个单行矩阵 $\boldsymbol{X}$,这个矩阵的列数就

是该空间的维数.在线性映射下向量的象由同样类型的矩阵 $\boldsymbol{XA}$ 表示,即是原向量的表示矩阵乘以该线性映射的表示矩阵之积.

线性形式关于基底的表示矩阵

令 y^* 是向量空间 F 上的线性形式.依定义,y^* 就是把 F 映入 K 的线性映射.若$(\boldsymbol{\varepsilon}_\mu)$ 是 F 的基底,$\boldsymbol{l}$ 是 K 的标准基度,则 y^* 的表示矩阵由如下公式得到

$$\begin{cases} y^*(\boldsymbol{\varepsilon}_1) = y_1^* \boldsymbol{l} \\ y^*(\boldsymbol{\varepsilon}_2) = y_2^* \boldsymbol{l} \\ \vdots \\ y^*(\boldsymbol{\varepsilon}_p) = y_p^* \boldsymbol{l} \end{cases}$$

即

$$\boldsymbol{Y}^* = \begin{pmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_p^* \end{pmatrix}$$

这里 $\boldsymbol{Y}^*$ 是 1 列矩阵.

令 f 是把 E 映入 F 的线性映射,映射 $y^* \circ f$ 就是把 E 映入 K 的线性映射,即 E 上的线性形式,亦即 E 的对偶空间 E^* 的元素.若 $\boldsymbol{A}$ 是 f 关于 E 的基底$(\boldsymbol{l}_i)$和 F 的基底$(\boldsymbol{\varepsilon}_\mu)$ 的表示矩阵,则 $y^* \circ f$ 关于$(\boldsymbol{l}_i)$ 与 $\boldsymbol{l}$ 的表示矩阵是 $\boldsymbol{AY}^*$.

所以,由 f 导出一个把 F^* 映入 E^* 的线性映射 f^*,其定义如下

$$f^*: y^* \in F^* \to y^* \circ f \in E^*$$

这是 f 的对偶映射.如果要得到这个映射在 F^* 和 E^* 关于已知基底的对偶基下的表示,就要考虑转置(交换行与列)矩阵,再按相应的形式处理.

3 线性算子与方阵

线性算子环与方阵环

a) 给了 n 维向量空间 E,使 E 映入自身的线性映射称为 E 的线性算子(或 E 的自同态映射).

只要假定 F 就是 E,前面的整个理论都可用于线性算子.为了用矩阵表示线性算子 f,只需给出 E 的基底$(\boldsymbol{l}_i)$,并将向量 $f(\boldsymbol{l}_i)$ 用这基底表出

$$f(\boldsymbol{l}_i) = \sum_{j=1}^{n} a_i^j \boldsymbol{l}_i$$

因此,算子 f 在基底$(\boldsymbol{l}_i)$(一般情形的基底$(\boldsymbol{\varepsilon}_\mu)$与$(\boldsymbol{l}_i)$一致)下由一个 $n \times n$ 矩阵(方阵)表示

$$\boldsymbol{A} = (a_i^j)$$





b) 设 $L(E,E)$ 是 E 的线性算子的集合；$L(E,E)$ 自然具有向量空间的结构.不仅如此，E 的两个线性算子的积 $g\circ f$ 恒有定义，仍为 E 的线性算子.在 $L(E,E)$ 中的这种合成法则是结合的，关于加法右分配且左分配，但不是交换的.最后，存在一个值得注意的线性算子，即恒等算子 e：与 E 的每个向量对应的是该向量自身；对任意 $f\in L(E,E)$，有

$$f\circ e = e\circ f = f$$

因此，上述合成法则有一单位元素.若 $\alpha\in K$，则

$$f\circ(\alpha e) = (\alpha e)\circ f = \alpha f$$

所以，用纯量 α 相乘，可视为线性算子乘积的特例(左乘或右乘以 αe 的积).

这样一来，在 $L(E,E)$ 上就有两个合成法则：

一个法则用加法表示，使 $L(E,E)$ 具有阿贝尔群的结构；另一个法则 $\circ$ 表示，满足非可换环的公理，有单位元素.

定理 线性算子的加法与乘法在 $L(E,E)$ 上定义一个有单位元素的非可换环结构.

c) 设 $M(n,n)$ 是 $n\times n$ 方阵的结合；两个这样的矩阵 $\boldsymbol{A},\boldsymbol{B}$ 的乘法恒有意义，$\boldsymbol{AB} = (\sum_j b_j^k a_i^j)$.在 E 中选定基底($\boldsymbol{e}_i$)，考虑把 $L(E,E)$ 映成 $M(n,n)$ 的一一映射 ϕ 相应于每个线性算子 f，映射 ϕ 给出这个算子的矩阵表示 $\boldsymbol{A} = \phi(f)$.根据矩阵加法与乘法的定义有

$$\phi(f+g) = \phi(f)+\phi(g),\phi(g\circ f) = \phi(g)\cdot\phi(f)$$

前面对 $L(E,F)$ 得到的结果对于 $M(n,n)$ 自然成立.特别，矩阵的乘积有单位元素，即恒等算子的表示矩阵

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_i^j)$$

由上可得如下定理：

定理 $n\times n$ 矩阵的加法与乘法在 $M(n,n)$ 上定义一非可换环结构，有单位元素并同构于 $L(E,E)$ 的环结构.

这样的环可以有零因子，即有可能两个矩阵的积为零，但其每一因子都不为零，如下例所示

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

最后，容易定义系数在 K 中，变元(向量)在算子环或方阵环中的多项式.令

$$\boldsymbol{A}^0 = \boldsymbol{I},\boldsymbol{A}^1 = \boldsymbol{A},\boldsymbol{A}^2 = \boldsymbol{AA},\cdots,\boldsymbol{A}^p = \boldsymbol{A}^{p-1}\boldsymbol{A}$$

立即有

$$A^pA^q = A^{p+q} = A^qA^p$$

单项式是 A 的幂与$\alpha \in K$的积αA^p. 多项式是有限个单项式的和. 对于系数与变元都在 K 中的任何多项式

$$P(\zeta) = \sum_{k=0}^{p} \alpha_k \zeta^k$$

我们可以相应地给出一个矩阵变元多项式

$$P(A) = \sum_{k=0}^{p} \alpha_k A^k \quad (A \in M(n,n))$$

若 K 是复数体,可取 $P(\zeta)$ 如下

$$P(\zeta) = \alpha_p \prod_{k=1}^{p} (\zeta - \mu_k)$$

可以立即验明

$$P(A) = \alpha_p \prod_{k=1}^{p} (A - \mu_k I)$$

正则线性算子与正则矩阵

a) E 的线性算子f叫做正则的,如果它是把 E 映成E 的线性映射. f是正则的充要条件是$f(E)$与 E 重合,即f的秩是n. 若f的秩为n,则f是一一的,并确定一把 E 映成E 的同构映射,我们把这样的 f 叫做自同构映射,非正则的算子叫奇异算子.

两个正则线性算子 f与g 的积$g \circ f$显然是正则的线性算子,因它把 E 映成 E. 反之,若 f 与 g 是两个线性算子,使得 $g \circ f$ 正则,则 f 与 g 全都是正则算子. 事实上

$$\mathrm{rang}(g \circ f) = \dim g(f(E)) \leqslant \dim f(E) = \mathrm{rang}(f)$$

由于 $\mathrm{rang}(g \circ f) = n$,有 $\mathrm{rang}\, f = n$,故 f 正则,$f(E) = E$;由于 $g(E)$ 的维数是 n,g 也正则.

设 f 正则,故 f 一一,所以对于每个 $y \in E$,相应地有唯一一个向量 x,使得 $y = f(x)$;把 E 映成E 的一一映射$y \to x$ 是f 的逆映射f^{-1} 使得对每一 $x \in E$ 有

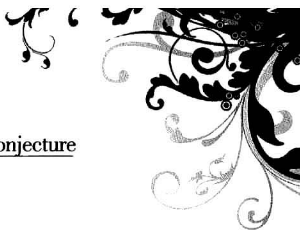
$$f^{-1}[f(x)] = x \tag{7}$$

这就是映射f^{-1}的特性. 按照与前述相同的论证可见,f^{-1}是线性的,所以每个正则线性算子 f 都有逆映射f^{-1},它也是正则线性算子,故也有逆算子. 显然,f^{-1} 的逆是 f. 关系(7) 以及交换 f 和f^{-1} 而得的关系可以改写为

$$f^{-1} \circ f = f \circ f^{-1} = e \tag{8}$$

若两个线性算子 f与g 使得$g \circ f = e$,则它们是正则的,并且对于每个 $x \in E$ 有





$$g(f(\boldsymbol{x})) = \boldsymbol{x}$$

这表明 g 是正则线性算子 f 的逆算子.于是下面定理成立:

定理 乘法 $\circ$ 使 E 的正则线性算子的集合具有群结构,这个群称为 E 的自同构群.

若 f 与 g 是两个正则线性算子,则 $g \circ f$ 正则,故存在逆算子,就是 $f^{-1} \circ g^{-1}$.这一性质对任何一个群都成立,因为

$$f^{-1} \circ g^{-1} \circ g \circ f = f^{-1} \circ f = e$$

b) 每个正则线性算子的表示矩阵叫做正则矩阵.要使一个矩阵是正则矩阵,充要条件是:它的秩是 n.选定 E 的一个基底,考虑映射 Φ:对于每个正则线性算子,相应地给出这个算子的矩阵表示.这个映射可以将前述的有关正则算子的结果转化成正则矩阵的结果.

特别,对每一 $n \times n$ 正则矩阵 $\boldsymbol{A}$,存在唯一 $n \times n$ 矩阵 $\boldsymbol{A}^{-1}$,使得对任意 $1 \times n$ 矩阵 $\boldsymbol{X}$

$$\boldsymbol{X}\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{X}$$

即

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$$

$\boldsymbol{A}^{-1}$ 是正则的,称为 $\boldsymbol{A}$ 的逆矩阵,还满足

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$$

上面的讨论表述成下面定理:

定理 $n \times n$ 矩阵的乘法使得由 K 中元素组成的 $n \times n$ 正则矩阵的集合具有群结构,这个群同构于 K 上的 n 维向量空间的自同构群,称为 K 的 n 元线性群,记为 $GL(n,K)$.

若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个 $n \times n$ 正则矩阵,则 $\boldsymbol{AB}$ 正则,并有逆 $\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.还要注意,若 $\boldsymbol{A}$ 正则,则等式 $\boldsymbol{AB} = \boldsymbol{O}$ 与 $\boldsymbol{BA} = \boldsymbol{O}$ 中任何一个都蕴涵 $\boldsymbol{B} = \boldsymbol{O}$.因为,设 $\boldsymbol{A}^{-1}$ 是 $\boldsymbol{A}$ 的逆,由 $\boldsymbol{AB} = \boldsymbol{O}$ 得 $\boldsymbol{A}^{-1}\boldsymbol{AB} = \boldsymbol{O}$,或 $\boldsymbol{B} = \boldsymbol{O}$,另一等式仿此可得.

行列式的作用

a) 每个 $n \times n$ 方阵都有其相应的行列式 $D(\boldsymbol{A})$(其值属于 K).行列式为零的充要条件是行向量组线性相关,即矩阵的秩小于 n.所以,一个矩阵是正则的充要条件是它的行列式 $\neq 0$.

若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是两个 $n \times n$ 矩阵,常用的行列式乘法如下

$$D(\boldsymbol{AB}) = D(\boldsymbol{A})D(\boldsymbol{B})$$

b) 方程组的克莱姆法则使我们直接得到正则矩阵 $\boldsymbol{A}$ 的逆矩阵 $\boldsymbol{A}^{-1}$ 的元素.

设 $\boldsymbol{Y}$ 是已知的 $n \times 1$ 矩阵,$\boldsymbol{X}$ 是 $n \times 1$ 未知矩阵,则矩阵方程

$$\boldsymbol{AX} = \boldsymbol{Y}$$

化为一组 n 个线性方程

$$\sum_{i=1}^{n} a_i^j x^i = y^j \quad (j = 1,2,\cdots,n)$$

这是克莱姆方程组.若 α_i^j 是 $D(A)$ 按 i 行或 j 列的展开式中元素 a_j^i 的系数,则

$$x^i = \frac{\sum_{j=1}^{n} \alpha_j^i y^j}{D(A)} = \sum_{j=1}^{n} \frac{\alpha_j^i}{D(A)} y^j = \sum_{j=1}^{n} b_j^i y^j$$

这里令

$$b_j^i = \frac{\alpha_j^i}{D(A)} \tag{9}$$

若 $\boldsymbol{B} = (b_j^i)$,则有下列矩阵关系

$$\boldsymbol{X} = \boldsymbol{BY}$$

$\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的逆矩阵.

利用方阵确定任意矩阵的秩

给了 $n \times p$ 矩阵 $\boldsymbol{A}$,$q \leqslant n, p$,从 $\boldsymbol{A}$ 中划去 $n - q$ 行与 $p - q$ 列得到的 $q \times q$ 方阵,称为从 $\boldsymbol{A}$ 选出的 q 阶方阵.矩阵 $\boldsymbol{A}$ 的秩可利用下列定理确定.

定理 矩阵 $\boldsymbol{A}$ 的秩等于从 $\boldsymbol{A}$ 中可以选出的正则方阵的最高阶数.

设 $\boldsymbol{A}$ 的秩为 r.要证明:不可能从 $\boldsymbol{A}$ 中选出一个正则方阵,其阶数 $s > r$,但却能从 $\boldsymbol{A}$ 中至少选出一个正则方阵,其阶数为 r.

因 $\boldsymbol{A}$ 的秩为 r,故 $\boldsymbol{A}$ 的 $s(> r)$ 个任意行向量是线性相关的.设 $\boldsymbol{B}$ 是从 $\boldsymbol{A}$ 选出的 s 阶方阵.它的行包含在 $\boldsymbol{A}$ 的 s 行中;所以,$\boldsymbol{B}$ 的行向量组是线性相关的,$\boldsymbol{B}$ 的秩不是 s,故 $\boldsymbol{B}$ 非正则.

定理的后一部分证明稍长.考虑向量空间 K^p,$(\boldsymbol{\varepsilon}_\mu)$ 是其标准基底.

若 $\boldsymbol{A}$ 的秩为 r,则 $\boldsymbol{A}$ 中存在一组 r 个自由行向量,为简化记号,假定就是 $\boldsymbol{A}$ 的前 r 行,记为 $(\boldsymbol{\lambda}_a)(a = 1,\cdots,r)$.令 V_1 是 K^p 中由 $(\boldsymbol{\lambda}_a)$ 生成的向量子空间.

若 $r < p$,则至少存在一 $\boldsymbol{\varepsilon}_\mu$ 不属于 V_1,所以同诸 $\boldsymbol{\lambda}_a$ 线性无关;将这一 $\boldsymbol{\varepsilon}_\mu$ 加到 $\boldsymbol{\lambda}_a$ 中,再考虑这样得到的一组新向量.容易看出,在 $(\boldsymbol{\varepsilon}_\mu)$ 中总是可以选取 $p - r$ 个向量,同诸 $\boldsymbol{\lambda}_a$ 一起成为 K^p 的一个基底;设这 $p - r$ 个向量是 $(\boldsymbol{\varepsilon}_{r+1},\cdots,\boldsymbol{\varepsilon}_p)$,或简记为 $(\boldsymbol{\varepsilon}_\lambda)(\lambda = r + 1,\cdots,p)$.

用 V_2 表示 K^p 中由 r 个向量 $(\boldsymbol{\varepsilon}_1,\cdots,\boldsymbol{\varepsilon}_r)$ 或 $(\boldsymbol{\varepsilon}_a)$ 生成的子空间,用 V_3 表示由 $(\boldsymbol{\varepsilon}_\lambda)$ 生成的子空间.显然,V_1 与 V_2 都是 V_3 的补子空间.令 $\boldsymbol{x} = \sum_{\mu=1}^{p} x^\mu \boldsymbol{\varepsilon}_\mu \in V_1$,$K^p$ 是 V_2 与 V_3 的直和,所以

$$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z} \quad (\boldsymbol{y} \in V_2, \boldsymbol{z} \in V_3)$$

$$\boldsymbol{y} = \sum_{a=1}^{r} x^a \boldsymbol{\varepsilon}_a, \boldsymbol{z} = \sum_{\lambda=r+1}^{p} x^\lambda \boldsymbol{\varepsilon}_\lambda$$

由 $\boldsymbol{x} \to \boldsymbol{y}$ 所定义的把 V_1 映入 V_2 的映射 f 显然是线性的.它是一一的,因为





只有对同时属于 V_1 与 V_3 的向量 $\boldsymbol{x}$,即零向量,才能 $\boldsymbol{y} = \boldsymbol{0}$. 因此 f 的秩等于 V_1 的维数 r,而 f 是把 V_1 映成 V_2 的同构映射.

在这同构映射下,与向量

$$\boldsymbol{\lambda}_a = \sum_{\mu=1}^{p} a_a^{\mu}\boldsymbol{\varepsilon}_{\mu} \quad (a = 1,\cdots,r)$$

对应的向量是

$$f(\boldsymbol{\lambda}_a) = \sum_{b=1}^{r} a_a^{b}\boldsymbol{\varepsilon}_b \quad (a = 1,\cdots,r)$$

在 V_1 的基底($\boldsymbol{\lambda}_a$) 和 V_2 的基底($\boldsymbol{\varepsilon}_b$) 下,同构映射 f 的表示矩阵 $\boldsymbol{B} = (a_a^b)$ 有秩 r,故为正则矩阵,这就是从 $\boldsymbol{A}$ 中选出的 r 阶正则矩阵.

反变与相似矩阵

a) 设 g 是 E 的正则线性算子,($\boldsymbol{l}_i$) 是 E 的基底,$\boldsymbol{P}$ 是 g 在此基底下的正则表示矩阵. n 个向量 $\boldsymbol{\varepsilon}_i = g(\boldsymbol{l}_i)$ 构成自由向量,它们确定 E 的另一基底. 反之,给了 E 的一个基底($\boldsymbol{\varepsilon}_i$),则存在唯一的线性算子 g 将基底($\boldsymbol{l}_i$) 映成基底($\boldsymbol{\varepsilon}_i$):这个算子显然是正则的,对于 $\boldsymbol{x}' = \sum x'^i\boldsymbol{l}_i$ 相应地给出 $g(\boldsymbol{x}') = \sum x'^i\boldsymbol{\varepsilon}_i$.

向量 $\boldsymbol{x} = g(\boldsymbol{x}')$ 是 E 的向量,在基底($\boldsymbol{\varepsilon}_i$) 下由 $1 \times n$ 矩阵确定

$$\boldsymbol{X}' = (x'^i)$$

在原来的基底($\boldsymbol{l}_i$) 下,它由下面矩阵确定

$$\boldsymbol{X} = \boldsymbol{X}'\boldsymbol{P}$$

因此,给了两个基底($\boldsymbol{l}_i$) 与($\boldsymbol{\varepsilon}_i$),如果在前一个基底下由正则矩阵 $\boldsymbol{P}$ 表示的线性算子把前一个基底变成后一个基底,那么在前一基底下由矩阵 $\boldsymbol{X}$ 表示的向量 $\boldsymbol{x}$ 在后一基底下则表为下面的矩阵

$$\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{P}^{-1} \tag{10}$$

这就是所谓向量分量的反变式变换的意思.

b) 设空间 E 有两个基底($\boldsymbol{l}_i$) 与($\boldsymbol{\varepsilon}_i$),$\boldsymbol{A}$ 是 $n \times n$ 矩阵,对于($\boldsymbol{l}_i$) 表示某一线性算子 f,如果 $\boldsymbol{X}$ 是向量 $\boldsymbol{x} \in E$ 在基底($\boldsymbol{l}_i$) 下的表示,则向量 $f(\boldsymbol{x})$ 的表示为

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{A}$$

关于($\boldsymbol{\varepsilon}_i$),$\boldsymbol{x}$ 与 $f(\boldsymbol{x})$ 分别表为

$$\boldsymbol{X}' = \boldsymbol{X}\boldsymbol{P}^{-1}, \boldsymbol{Y}' = \boldsymbol{Y}\boldsymbol{P}^{-1}$$

因此

$$\boldsymbol{Y}' = \boldsymbol{X}'\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}$$

f 关于基底($\boldsymbol{\varepsilon}_i$) 的表示矩阵是

$$\boldsymbol{B} = \boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1} \quad (\boldsymbol{P} \in GL(n,K))$$

矩阵 $\boldsymbol{B}$ 叫做 $\boldsymbol{A}$ 关于线性群的相似矩阵;$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是同一线性算子在不同基底下的表示矩阵. 特别,如果给出指标 i 所成序列的置换 π,则矩阵 $\boldsymbol{A} = (a_i^j)$ 与 $\boldsymbol{B} = (a_{\pi(j)}^{\pi(i)})$ 相似,因为它们是同一个线性算子关于基底($\boldsymbol{l}_i$) 与($\boldsymbol{l}_{\pi(j)}$) 的表示.

这时,我们说 $\boldsymbol{B}$ 由 $\boldsymbol{A}$ 置换而得.

容易看出相似是 $M(n,n)$ 的元素之间的等价关系:特别,传递性是明显的;若 $\boldsymbol{C}=\boldsymbol{QBQ}^{-1}(\boldsymbol{Q}\in GL(n,K))$,$\boldsymbol{B}=\boldsymbol{PAP}^{-1}$,则

$$\boldsymbol{C}=\boldsymbol{QPAP}^{-1}\boldsymbol{Q}^{-1}=(\boldsymbol{QP})\boldsymbol{A}(\boldsymbol{QP})^{-1}$$

这里 $\boldsymbol{QP}$ 是正则矩阵.

4 谱的概念,方程关于线性群的法式

法式问题

相似既然是 $M(n,n)$ 中的等价关系,让我们考虑与一个已知矩阵相似的矩阵的集合,这个集元的元素彼此相似,称为关于线性群的等价类,于是 $M(n,n)$ 被分成等价类.

知道了等价类中的一个矩阵,这等价类就确定了,因之自然想在等价类的诸矩阵中找出一个尽可能简单的矩阵,使得对于每一等价类可以相应地唯一决定(精确到置换)其中一个矩阵,说明该等价类的特性.这个矩阵与等价类中各矩阵都相似,叫做该等价类中任何矩阵 $\boldsymbol{A}$ 的法式.

矩阵化为法式的问题涉及矩阵 $\boldsymbol{A}$ 的元素的某些函数,这些函数有一个重要性质:对所有与 $\boldsymbol{A}$ 相似的矩阵取同一值.这样的函数称为矩阵 $\boldsymbol{A}$ 关于线性群的不变量.$n\times n$ 矩阵关于线性群的法式问题是一个相当困难的代数问题;我们只能在下节讨论一个特别情形.

对角矩阵

最简单的一种方阵是对角方阵:$n\times n$ 矩阵$\boldsymbol{A}$ 叫做对角矩阵,如果它的元素除去主对角线上的外都为0;例如单位矩阵 $\boldsymbol{I}$ 就是对角矩阵.设 α_i 是对角矩阵$\boldsymbol{A}$ 的对角线上的第 i 个元素.若$(\boldsymbol{l}_i)$ 是 E 的一个基底,则 $\boldsymbol{A}$ 表示一线性算子f,满足

$$f(\boldsymbol{l}_i)=\alpha_i\boldsymbol{l}_i$$

因此,E 中由 $\boldsymbol{l}_i$ 生成的每个一维子空间对f是不变的.记 $\boldsymbol{A}=(\alpha_i)$.

所有对角矩阵构成 $M(n,n)$ 的可换子环.事实上,若 $\boldsymbol{A}=(\alpha_i)$,$\boldsymbol{B}=(\beta_i)$ 是两个对角矩阵,它们对于$(\boldsymbol{l}_i)$ 表示两个算子 f 与 g,使得

$$f(\boldsymbol{l}_i)=\alpha_i\boldsymbol{l}_i,\ g(\boldsymbol{l}_i)=\beta_i\boldsymbol{l}_i$$

所以

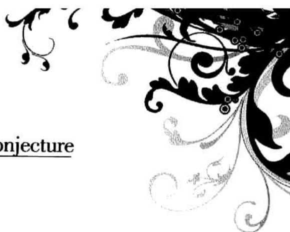
$$f(\boldsymbol{l}_i)+g(\boldsymbol{l}_i)=(\alpha_i+\beta_i)\boldsymbol{l}_i$$

$$g(f(\boldsymbol{l}_i))=\beta_i\alpha_i\boldsymbol{l}_i$$

$\boldsymbol{A}+\boldsymbol{B}=(\alpha_i+\beta_i)$,$\boldsymbol{BA}=(\beta_i\alpha_i)$ 是对角矩阵,所以积 $\boldsymbol{BA}$ 显然是可换的.

我们自然要研究对角矩阵是否能成为法式的问题.在一般情形下,这是不





可能的,不过下节要提出的一个重要特例却是可能的.

特征向量与特征值

以后我们都假定 K 是复数体 $\mathbf{C}$.

a) 设 E 是 $\mathbf{C}$ 上的向量空间.给了 E 的线性算子 f,我们来求对 f 不变的一维子空间.若 $\boldsymbol{P} \neq \mathbf{0}$ 属于这样的子空间,则存在一纯量 χ 满足

$$f(\boldsymbol{P}) = \chi \boldsymbol{P} \quad \chi \in \mathbf{C} \tag{11}$$

上式可写成

$$(\chi e - f)(\boldsymbol{P}) = \mathbf{0} \tag{12}$$

若存在一个这样的向量 $\boldsymbol{P}$,则 $\chi e - f$ 非正则.反之,若 $\chi e - f$ 非正则(奇异),则 $(\chi e - f)^{-1}(0)$ 不是 0,即存在 $\boldsymbol{P} \neq \mathbf{0}$ 满足(11)或(12).由此得到下述定义:

定义 使得 $\chi e - f$ 为奇异算子的每一 $\chi \in \mathbf{C}$ 称为线性算子 f 的特征值.使得 $f(\boldsymbol{P}) = \chi \boldsymbol{P}$ 的任何向量 $\boldsymbol{P} \neq \mathbf{0}$ 称为属于特征值 χ 的特征向量.f 的特征值的集合称为 f 的谱.

b) 在 E 内取一个基底 $(\boldsymbol{l}_i)$,设 $\boldsymbol{A}$ 是 f 的表示矩阵,则 $(\chi e - f)$ 的表示矩阵为 $(\chi \boldsymbol{I} - \boldsymbol{A})$.要使 χ 是 f 的特征值,充要条件是 $D(\chi \boldsymbol{I} - \boldsymbol{A}) = 0$,即 χ 是下式的零点

$$\psi(\chi) = D(\chi \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \chi - a_1^1 & -a_1^2 & \cdots & -a_1^n \\ -a_2^1 & \chi - a_2^2 & \cdots & -a_2^n \\ \vdots & \vdots & & \vdots \\ -a_n^1 & -a_n^2 & \cdots & \chi - a_n^n \end{vmatrix} \tag{13}$$

要注意,按定义,对于两个相似矩阵,它们的 n 阶多项式 $\psi(\chi)$ 是一样的.事实上,若 $\boldsymbol{B} = \boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1}$,则

$$\chi \boldsymbol{I} - \boldsymbol{B} = \boldsymbol{P}(\chi \boldsymbol{I} - \boldsymbol{A})\boldsymbol{P}^{-1}$$
$$D(\chi \boldsymbol{I} - \boldsymbol{B}) = D(\chi \boldsymbol{I} - \boldsymbol{A})$$

我们把多项式 $\psi(\chi)$ 称为算子 f 或矩阵 $\boldsymbol{A}$ 的特征多项式.其零点 $\chi_1, \chi_2, \cdots, \chi_n$ 可同可不同,它们就是 f 的特征值,也叫 $\boldsymbol{A}$ 的特征值.$\psi(\chi)$ 的各系数都是 $\boldsymbol{A}$ 关于线性群的不变量.特别,$\psi(\chi)$ 的前两项显然如下

$$\psi(\chi) = \chi^n - \left(\sum_i a_i^i\right)\chi^{n-1} + \cdots$$

所以,函数

$$\operatorname{tr} \boldsymbol{A} = \sum a_i^i = \chi_1 + \chi_2 + \cdots + \chi_n$$

是 $\boldsymbol{A}$ 的不变量,称为矩阵 $\boldsymbol{A}$(或它在基底 $\boldsymbol{l}_i$ 下所表示的算子 f)的迹.

把具有互异特征值的矩阵化为法式

设 $\boldsymbol{A}$ 是元素在 C 中的 $n \times n$ 矩阵,它的特征值 $\chi_1, \chi_2, \cdots, \chi_n$ 互异.$(\boldsymbol{l}_i)$ 是 C

上的向量空间 E 的一个基底,设 f 是$\boldsymbol{A}$ 所表示的线性算子,以(χ_i)为特征值.我们来研究这个算子.对于每一特征值 χ_i,至少有一个相应的特征向量 $\boldsymbol{P}_i \neq \boldsymbol{0}$

$$f(\boldsymbol{P}_i) = \chi_i \boldsymbol{P}_i \quad (i = 1,2,\cdots,n) \tag{14}$$

如上得到的 n 个特征向量$\boldsymbol{P}_i$,当诸 χ_i 各不相同时构成一组自由向量.事实上,假定只有 $q(< n)$ 个向量 $\boldsymbol{P}_i$ 线性无关,例如前 q 个,则 $\boldsymbol{P}_n$ 可唯一地表成 $\boldsymbol{P}_\mu(\mu = 1,\cdots,q)$ 的线性组合

$$\boldsymbol{P}_n = \sum_{\mu=1}^{q} \alpha_\mu \boldsymbol{P}_\mu \quad (\alpha_\mu \text{ 不全为 } 0) \tag{15}$$

由此推出

$$f(\boldsymbol{P}_n) = \sum_\mu \alpha_\mu f(\boldsymbol{P}_\mu)$$

即

$$\chi_n \boldsymbol{P}_n = \sum_\mu \alpha_\mu \chi_\mu \boldsymbol{P}_\mu$$

若 $\chi_n = 0$,则诸 χ_μ 全不为0,故诸 $\boldsymbol{P}_\mu$ 有一个线性组合为0,但系数不全为0.因此 $\chi_n \neq 0$

$$\boldsymbol{P}_n = \sum_\mu \alpha_\mu \frac{\chi_\mu}{\chi_n} \boldsymbol{P}_\mu \tag{16}$$

此式应与(15)一致;若 $\alpha_\mu \neq 0$,则 $\chi_\mu = \chi_n$,故特征值不是互异的.

自由向量组$(\boldsymbol{P}_i)$构成 E 的一个基底.根据(14),在这一个基底下,算子 f 由如下对角矩阵表示

$$R(\boldsymbol{A}) = \begin{pmatrix} \chi_1 & 0 & \cdots & 0 \\ 0 & \chi_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \chi_n \end{pmatrix}$$

它与 $\boldsymbol{A}$ 相似.

因此,任何具有互异特征值的矩阵 $\boldsymbol{A}$ 与对角矩阵相似,对角线上的元素是 $\boldsymbol{A}$ 的特征值.在诸特征值非互异的情形,法式有更复杂的形式(约当法式).

要使矩阵 $\boldsymbol{B}$ 相似于矩阵$\boldsymbol{A}$ 必须它们有同样的特征值.如果诸特征值互异,这条件还是充分的,因为两个矩阵都相似于 $R(\boldsymbol{A})$.

方阵 $\boldsymbol{A}$ 的特征多项式$\psi(x)$有一个重要性质,我们来给出一个初等证明.

设 $\alpha_j^i(\chi)$ 是 $D(\chi\boldsymbol{I} - \boldsymbol{A})$ 按 i 行或j 列的展开式中i 行j 列元素的系数.这里 χ 的 $n-1$ 次多项式.令 $C(\chi) = (\alpha_j^i(\chi))$,则有

$$(\chi\boldsymbol{I} - \boldsymbol{A})^{-1} = \frac{C(\chi)}{\psi(\chi)}$$

所以

$$C(\chi)(\chi\boldsymbol{I} - \boldsymbol{A}) = \psi(\chi)\boldsymbol{I} \tag{17}$$



这里 $C(\chi)$ 是以矩阵为系数的 χ 的多项式

$$C(\chi) = \Gamma_1\chi^{n-1} + \Gamma_2\chi^{n-2} + \cdots + \Gamma_n \quad (\Gamma_i \in M(n,n))$$

若将 $\psi(\chi)$ 写成

$$\psi(\chi) = \chi^n + c_1\chi^{n-1} + \cdots + c_n$$

则由(17),比较系数得

$$\begin{aligned} \Gamma_1 &= 1 \\ \Gamma_2 - \Gamma_1\boldsymbol{A} &= c_1\boldsymbol{I} \\ &\vdots \\ \Gamma_n - \Gamma_{n-1}\boldsymbol{A} &= c_{n-1}\boldsymbol{I} \\ -\Gamma_n\boldsymbol{A} &= c_n\boldsymbol{I} \end{aligned}$$

依次用 $\boldsymbol{A}^n,\boldsymbol{A}^{n-1},\cdots,\boldsymbol{A}^0$ 乘上面的各式,然后各边相加便得

$$\psi(\boldsymbol{A}) = 0$$

哈密顿 - 凯莱定理 若 $\psi(\chi)$ 是方阵 $\boldsymbol{A}$ 的特征多项式,则

$$\psi(\boldsymbol{A}) = 0$$

第五讲　二次形式与厄米形式①

1　引　言

n 个实变量 $x_1, x_2, \cdots, x_n$ 的二次形式是这些变量的二次齐次多项式,其系数为实数,即

$$F = \sum_{i,j} a_{ij}x_ix_j \tag{1}$$

式(1)求和是关于脚标 i 和 j 各自独立地从 1 变到 n 而言;乘积 $x_ix_j = x_jx_i$ 是可换的,所以设 $a_{ij} = a_{ji}$,亦即 $n \times n$ 矩阵 $\boldsymbol{A} = (a_{ij})$ 是对称矩阵.式(1)还可写成

$$F = \sum_i a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j \tag{2}$$

二次形式的研究间接地属于"数学专修班"②的教学大纲;二次曲面方程实际上可以用齐次坐标 $x_1 = X, x_2 = Y, x_3 = Z, x_4 = T$ 写成 $F = 0$,即是

$$F = AX^2 + A'Y^2 + A''Z^2 + 2BYZ + \cdots + 2CT + \cdots + DT^2 = 0$$

形式(1)关于正交群的化简问题涉及"s 的方程"③.某些学校的教学大纲中明确提出二次形式,这可以对本讲演的观点加以修正.我们将直截了当地使用一些基本概念(线性变换,正交群,极形式的对称 $\boldsymbol{y} \cdot \boldsymbol{Ax} = \boldsymbol{Ay} \cdot \boldsymbol{x}$,等.这里 $\boldsymbol{Ax}$ 是向量 $\boldsymbol{x}$ 经过对称矩阵作用后的变形),以及其他几个讲演里已有的一些概念.我们将得到最初等的证明,好处是这些证明对任意多个变量都成立.对于迄今只注意研究二次曲面的教学工作,还有一个理由要加以修改,即是使教学工作更接近于绝大多数学生在物理和力学中将会碰到的那些应用.二次曲面可以很方便地用来表现二次形式,例如刚体的惯性椭球,所以一方面我们为二次曲面的讨论留有一席之地,同时我们着重研究线性变换

$$y_i = \sum_j a_{ij}x_j, a_{ij} = a_{ji} \tag{3}$$

即是 $\boldsymbol{y} = \boldsymbol{Ax}$($\boldsymbol{A} = (a_{ij})$ 是对称矩阵),以及变换(3)的特征值与特征向量.特征

① 原作者 P. Lelong(索尔本大学教授).
② 原文为 Mathématique Spécials,指法国中学里专门为投考理工学院所设的班.
③ 指下面的特征方程.



值的概念涉及许多应用(例如刚体的运动矩与旋转矩之间的对应关系;梁的弯曲力偶与中性线之间的对应关系,等).

本文在讨论形式(1)的同时,还要讨论所谓的厄米形式

$$\phi = \sum_{p,q} c_{pq}\bar{x}_p x_q \tag{4}$$

其系数 c_{pq} 是复数;a 与 $\bar{a}$ 表示两个共轭复数.形式 ϕ 关于两组变量 x_p,$\bar{x}_p$ 是双线性的.如果对于任何共轭复数 x_p,$\bar{x}_p$,ϕ 的值是实数,它就叫厄米形式.为此,必须且只需

$$c_{pq} = \bar{c}_{qp} \tag{5}$$

这是明显的(除了标号为 p,q 的变量外,让其余变量全为零).厄米形式有重要的作用,特别是在量子力学中.

对厄米形式的研究,即是它关于酉群的化简问题,要用到它的双线性特征.二次形式的研究只不过是一个特殊情形.事实上,如果在(4)中,c_{pq} 是实数,则由式(5)得 $c_{pq} = c_{qp}$.所以二次形式(1)是就实变量来讨论的实系数厄米形式.我们在研究厄米形式的同时,也顺便指出二次形式中可以进行的某些简化,这主要在于实数及其共轭并无区别,在各种特殊情况下将指出表达方式的修改,据此可以独立讨论实坐标空间 E_n 中的二次形式.

2 赋范空间 C_n 与 E_n

我们用 C_n 表示复数体上的 n 维向量空间.存在 C_n 的 n 个元素(或向量)$\boldsymbol{e}_i$ 构成 C_n 的一个基底 R_0,则诸 $\boldsymbol{e}_i$ 可将 C_n 的任一元素 $\boldsymbol{x}$ 表成如下形式

$$\boldsymbol{x} = \sum x_i \boldsymbol{e}_i$$

诸 x_i 是 n 个复数,称为 $\boldsymbol{x}$ 关于基底 $\boldsymbol{R}_0$ 的坐标.C_n 的自同态是把 C_n 映入自身的变换 $\boldsymbol{x}' = g(\boldsymbol{x})$,它保持向量空间的结构①.如果知道

$$\boldsymbol{e}'_i = g(\boldsymbol{e}_i) = \sum_j a_{ji}\boldsymbol{e}_j \tag{6}$$

g 就确定了.这时 $\boldsymbol{x}' = \sum x_i g(\boldsymbol{e}_i) = \sum_{i,j} x_i a_{ji}\boldsymbol{e}_j$.在基底 $\boldsymbol{R}_0$ 下,变换 $\boldsymbol{x}' = g(\boldsymbol{x})$ 表成

$$x'_i = \sum_j a_{ij}x_j \tag{7}$$

因此我们说自同态 $\boldsymbol{x}' = g(\boldsymbol{x})$,即是(6),在基底 $\boldsymbol{R}_0$ 下由矩阵 $\boldsymbol{A} = (a_{ij})$ 表示;实际上,这个矩阵使我们可以在基底 $\boldsymbol{R}_0$ 下把 $\boldsymbol{x}'$ 的坐标按公式(7)表成 $\boldsymbol{x}$ 的坐标的函数.在(7)中视向量 $\boldsymbol{x} = (x_1,\cdots,x_n)$ 为 $1 \times n$ 矩阵便可将这些坐标缩写

① 参考 G. Choquet 的讲演:向量空间.

为

$$x' = Ax \tag{8}$$

依次施行两个自同态 $x'' = Bx'$，$x' = Ax$ 便确定一自同态 $x'' = BAx = Cx$，这里 $C = BA$ 是矩阵积

$$c_{ij} = \sum_s b_{is} a_{sj}$$

若 A 表示矩阵 (a_{ij})，我们将元素为

$$a'_{ji} = a_{ij}$$

的矩阵 A' 称为 A 的转置.

显然有：$(A')' = A$，$(BA)' = A'B'$.

C_n 中的纯量积与范数　两个向量 a 与 b 的厄米纯量积是指一个复数 $a \cdot b$，定义为向量 a 与 b 的函数，具有下列性质：

a) $a \cdot b = \overline{b \cdot a}$（非可换）；

b) $a \cdot (\lambda b_1 + \mu b_2) = \lambda(a \cdot b_1) + \mu(a \cdot b_2)$，$(\lambda a_1 + \mu a_2) \cdot b = \overline{\lambda}(a_1 \cdot b) + \overline{\mu}(a_2 \cdot b)$，$\lambda$ 与 μ 是任意复数，$\overline{\lambda}$ 与 $\overline{\mu}$ 是它们的共轭复数；

c) $a \cdot a$ 是一非负实数，仅当 a 是零向量时为零. a 的范数是指非负实数

$$\| a \| = \sqrt{a \cdot a}$$

若将 a，b 在一组基底 $R_0 = (e_i)$ 下表出，则有

$$a \cdot b = \left(\sum_i a_i e_i\right) \cdot \left(\sum_j b_j e_j\right) = \sum_{i,j} \overline{a}_i b_j (e_i \cdot e_j)$$

如果我们确定了纯量积 $(e_i \cdot e_j)$，那么厄米纯量积便完全确定了.

两个向量 a 与 b 叫做正交的，如果 $a \cdot b = 0$. 范数 $\| a \| = 1$ 的向量 a 叫做单位向量. 一组两两正交的单位向量叫做法正交的. 一个基底 R_0 叫做法正交基底，如果它由法正交向量组成.

定理 1　若 $a_1, a_2, \cdots, a_m$ 是 m 个线性无关的向量，则存在一组法正交向量 $b_1, b_2, \cdots, b_m$ 使得对每个整数 $k(1 \leqslant k \leqslant m)$，向量组 $(a_1, \cdots, a_k)$ 与 $(b_1, \cdots, b_k)$ 都是同一向量子空间的两个基底.

对于 $m = 1$，定理显然，因定理的假设蕴涵 $\| a_1 \| \neq 0$，故只需取 $b_1 = \dfrac{a_1}{\| a_1 \|}$ 即可.

关于 m 用归纳法，设结论对 $m - 1$ 真：故存在 $m - 1$ 个向量 $b_1, \cdots, b_{m-1}$ 使得对于 $k \leqslant m - 1$，$(a_1, \cdots, a_k)$ 与 $(b_1, \cdots, b_k)$ 是同一向量子空间的基底，特别，对 $k = m - 1$ 是如此.

令

$$c = a_m - \sum_i b_i \overline{(b_i \cdot a_m)}$$

显然 c 不属于由 $(a_1, \cdots, a_{m-1})$ 生成的向量空间，否则 a_m 也是这样，此与假设不





合.

定义

$$\boldsymbol{b}_m = \boldsymbol{c}(\|\boldsymbol{c}\|^{-1})$$

于是有

$$\|\boldsymbol{b}_m\| = 1, \boldsymbol{b}_m \cdot \boldsymbol{b}_k = \|\boldsymbol{c}\|^{-1}[\boldsymbol{a}_m \cdot \boldsymbol{b}_k - \overline{(\boldsymbol{b}_k \cdot \boldsymbol{a}_m)}] = 0$$

故 $\boldsymbol{b}_1, \cdots, \boldsymbol{b}_m$ 是法正交组,与 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 生成同一个向量空间.

注 1) 厄米积 $\boldsymbol{a} \cdot \boldsymbol{b}$ 不是可换的,但满足 a).

2) p 个两两正交的非零向量 $\boldsymbol{c}_1, \cdots, \boldsymbol{c}_p$ 线性无关:事实上,从 $\sum\limits_1^p \lambda_i \boldsymbol{c}_i = 0$ 推出

$$\boldsymbol{c}_i \cdot \sum \lambda_j \boldsymbol{c}_j = \lambda_i \|\boldsymbol{c}_i\|^2 = 0, \text{故 } \lambda_i = 0$$

上面定理的证明给出了依次构造向量 $\boldsymbol{b}_i$ 的方法.首先正交化向量组 $\boldsymbol{a}_i$,即构造向量 $\boldsymbol{c}_i$ 如下

$$\begin{gathered}
\boldsymbol{c}_1 = \boldsymbol{a}_1 \\
\boldsymbol{c}_2 = \boldsymbol{a}_2 + \lambda_2^1 \boldsymbol{c}_1 \\
\boldsymbol{c}_3 = \boldsymbol{a}_3 + \lambda_3^1 \boldsymbol{c}_1 + \lambda_3^2 \boldsymbol{c}_2 \\
\vdots \\
\boldsymbol{c}_n = \boldsymbol{a}_n + \lambda_n^1 \boldsymbol{c}_1 + \cdots + \lambda_n^{n-1} \boldsymbol{c}_{n-1}
\end{gathered}$$

对任意的 λ_p^q,每一个向量 $\boldsymbol{c}_i$ 不为零.依次写出 $\boldsymbol{c}_2 \cdot \boldsymbol{c}_1 = 0, \cdots, \boldsymbol{c}_m \cdot \boldsymbol{c}_q = 0$,就可以确定 $\lambda_2^1, \cdots, \lambda_m^q$,即是 $\lambda_2^1 \|\boldsymbol{c}_1\|^2 + \boldsymbol{a}_2 \cdot \boldsymbol{c}_1 = 0$ 定出 λ_2^1,然后

$$\lambda_3^1 \|\boldsymbol{c}_1\|^2 + \boldsymbol{a}_3 \cdot \boldsymbol{c}_1 = 0 \text{ 与 } \lambda_3^2 \|\boldsymbol{c}_2\|^2 + \boldsymbol{a}_3 \cdot \boldsymbol{c}_2 = 0$$

定出 λ_3^1 与 λ_3^2,等.其次再将向量组($\boldsymbol{c}_i$)法化(变成单位向量),为此令 $\boldsymbol{b}_i = \dfrac{\boldsymbol{c}_i}{\|\boldsymbol{c}_i\|}$;这样诸 $\boldsymbol{b}_i$ 就是所求的法正交组.

E_n 中的实纯量积 用 E_n 表示实数体上的 n 维向量空间.取定 E_n 中一组 n 个线性无关向量后,E_n 的元素或向量 $\boldsymbol{a}$ 就由其实数坐标(a_i)确定.前述理论中唯一要修改的只是纯量积的定义.纯量积总是记为 $\boldsymbol{a} \cdot \boldsymbol{b}$,定义为 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的实函数,满足:

a) $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$;

b) $\boldsymbol{a} \cdot (\lambda \boldsymbol{b}_1 + \mu \boldsymbol{b}_2) = \lambda \boldsymbol{a} \cdot \boldsymbol{b}_1 + \mu \boldsymbol{a} \cdot \boldsymbol{b}_2$,$\lambda, \mu$ 是任意实数;

c) $\boldsymbol{a} \cdot \boldsymbol{a}$ 是正数,仅当 $\boldsymbol{a}$ 为零向量时为零.

当 $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ 时,也说 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 正交;当 $\|\boldsymbol{a}\| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}} = 1$ 时,也说 $\boldsymbol{a}$ 是单位向量;仍可如前面一样定义法正交基,同法证明定理 1,不过这时不必区别 $\boldsymbol{b}_m \cdot \boldsymbol{b}_k$ 与 $\boldsymbol{b}_k \cdot \boldsymbol{b}_m$,因实纯量积是可换的.

(C_n 中)厄米纯量积与(E_n 中)实纯量积的性质

a) 在向量 $\boldsymbol{e}_i$ 组成的法正交基 $\boldsymbol{R}_0$ 下将 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 表出,则

$$\boldsymbol{a}\cdot\boldsymbol{b}=\sum_{i,j}a_ib_j(\boldsymbol{e}_i\cdot\boldsymbol{e}_j)=\sum_i\bar{a}_ib_i,\ \|\boldsymbol{a}\|^2=\sum_i|a_i|^2$$

故有
$$|\boldsymbol{a}\cdot\boldsymbol{b}|=|\sum_i\bar{a}_ib_i|\leqslant\sum_i|a_i||b_i|$$

与
$$(\sum_i|\bar{a}_i||b_i|)^2\leqslant\sum_i|a_i|^2\cdot\sum_i|b_i|^2$$

后一不等式是由于 λ 的二次三项式

$$M(\lambda)=\sum_i(|a_i|+\lambda|b_i|)^2\geqslant 0$$

所以判别式非负,因此

$$|\boldsymbol{a}\cdot\boldsymbol{b}|\leqslant\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|\tag{9}$$

b) 由上推出

$$\|\boldsymbol{a}+\boldsymbol{b}\|^2=(\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}+\boldsymbol{b})=\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2+\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{b}\cdot\boldsymbol{a}$$

$$\|\boldsymbol{a}+\boldsymbol{b}\|^2\leqslant\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2+2\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|=(\|\boldsymbol{a}\|+\|\boldsymbol{b}\|)^2$$

所以

$$\|\boldsymbol{a}+\boldsymbol{b}\|\leqslant\|\boldsymbol{a}\|+\|\boldsymbol{b}\|\tag{10}$$

式(10) 叫三角形不等式,其证明既适用于 C_n 中的厄米纯量积,也适用于 E_n 中的纯量积;不论哪种情形,式(10) 可解释成欧氏几何中的一条定理:三角形中一边的长小于或等于其他两边的和.

相伴算子　设 $\boldsymbol{T}$ 是之前考察过的算子,即由矩阵(t_{ij})表示的自同态.相应于 $\boldsymbol{T}$ 有一个算子 $\boldsymbol{T}^*$,称为 $\boldsymbol{T}$ 的相伴算子,定义为

$$\boldsymbol{T}^*\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{a}\cdot\boldsymbol{T}\boldsymbol{b}\tag{11}$$

易见 $\boldsymbol{T}^*$ 是 C_n 的自同态.若 $\boldsymbol{T}^*$ 的表示矩阵为(t_{ij}^*),在法正交基底 $\boldsymbol{R}_0$ 下将式(11) 表出,则得

$$\sum_{i,j}t_{ij}^*\bar{a}_jb_i=\sum_{i,j}\bar{a}_jt_{ji}b_i$$

故关于基底 $\boldsymbol{R}_0$,$\boldsymbol{T}^*$ 的表示矩阵为

$$(t_{ij}^*)=(\bar{t}_{ji})\tag{12}$$

由式(12) 确定的矩阵(t_{ij}^*)称为矩阵(t_{ij})的相伴矩阵这是对 $\boldsymbol{T}=(t_{ij})$ 取转置矩阵 $\boldsymbol{T}'=(t_{ji})$,再取共轭矩阵而得,此外转置与共轭这两个运算是可换的.$\boldsymbol{T}$ 的相伴算子 $\boldsymbol{T}^*$ 的定义可以用初等方法提出:先由式(12) 定义相伴矩阵,然后验证等式(11).然而,从式(11) 出发直接定义 $\boldsymbol{T}^*$ 具有一般性,可以推广用来定义具有厄米纯量积的向量空间中任意算子的相伴算子.

不论从(11) 出发或(12) 出发,都可以证明相伴矩阵的下述性质:

a) $(\boldsymbol{A}^*)^*=.\boldsymbol{A}$.

b) $(\boldsymbol{B}\boldsymbol{A})^*=\boldsymbol{A}^*\boldsymbol{B}^*$.





c) 若 A 是 $(n \times n)$ 正则矩阵,即存在 A 的逆(或者 A 的行列式非零),使得 $AA^{-1} = I$,或者利用 b):$(A^{-1})^* A^* = I$;由此 $(A^{-1})^* = (A^*)^{-1}$.

厄米算子 现在来重点讨论成为自己的相伴算子的那些算子.算子 T 称为厄米算子(在空间 C_n 中),如果

$$y \cdot Tx = Ty \cdot x$$

也就是 $T = T^*$.量子力学的算子一般就是这类算子.我们将研究在什么条件下,由矩阵 $A = (a_{ij})$ 所确定的自同态 $Tx = Ax$ 是厄米算子;因为

$$y \cdot Ax = Ay \cdot x$$

或者由(12),$a_{ij} = \overline{a}_{ji}$,所以 $A = (a_{ij})$ 是厄米的,必须且只需下列形式

$$\phi(x) = x \cdot Ax = \sum_i (\overline{x}_i \sum_j a_{ij}x_j) = \sum_{i,j} a_{ij}\overline{x}_i x_j$$

是引言中所说的厄米形式.

E_n 中的对称算子 在空间 E_n 中,坐标是实数,纯量积是实数并且可换,我们说算子 T' 是算子,T 的转置算子,如果

$$T'a \cdot b = a \cdot Tb \tag{11)'$$

若 T 是 E_n 的自同态,则 T' 也是,T' 的表示矩阵 (t'_{ij}) 由下式给出

$$(t'_{ij}) = (t_{ji}) \tag{12)'$$

这是 $A = (t_{ij})$ 的转置矩阵.此外,还有 $A' = A^*$,这是因为矩阵的元素都是实数,所以其相伴矩阵与转置矩阵一致,算子 T 叫对称的,如果它与其转置重合,在法正交基底 R_0 下,由 $y = Ax$,$A = (a_{ij})$ 所表示的算子 Tx 是对称的,必须且只需 $a_{ij} = a_{ji}$,即矩阵 A 对称.形式

$$F(x) = x \cdot Ax = \sum a_{ij}x_j x_j \tag{13}$$

叫 R_0 下该算子的附属形式;两个向量 x 与 y 的函数 $F_1(x, y) = y \cdot Ax = Ay \cdot x$ 称为极形式,它也可写成 $x \cdot Ay$,因实纯量积是可换的.

特征向量与特征值[①] 设 T 是(C_n 中)厄米算子或者(E_n 中)对称算子.非零向量 x 使得

$$Tx = sx \tag{14}$$

时,称为 T 的特征向量,数 s 叫做该算子的特征值,x 叫属于 s 的特征向量.用 A 表示在基底 R_0 下 T 的表示矩阵.

特征值与特征向量的性质:

1) 存在 n 个特征值(实的或互异).事实上,命 I 是单位矩阵,(14) 可以写成 $(A - sI)x = 0$.让左端的 n 个坐标为零,便得 n 个线性方程;仅当

---

① 见 A. Lichnerowicz 的讲演:线性映射与矩阵.

$$D_A(s)=\begin{vmatrix} a_{11}-s & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-s & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}-s \end{vmatrix}=0$$

时,方程组才有解 $\boldsymbol{x}\neq\boldsymbol{0}$. 由于 $D_{\boldsymbol{A}}(s)=(-1)^n s^n+\cdots$,故至多存在 n 个特征值(相同或不同). 反之,若 s 是 $D_{\boldsymbol{A}}(s)=0$ 之根,则存在一向量 $\boldsymbol{x}\neq\boldsymbol{0}$ 满足 $\boldsymbol{A}\boldsymbol{x}=s\boldsymbol{x}$,所以是属于 s 的特征向量.

2) 依定义,特征向量与特征值仅与算子

$$\boldsymbol{y}=\boldsymbol{T}\boldsymbol{x} \tag{15}$$

有关,而与在基底 $\boldsymbol{R}_0$ 下该算子的矩阵表示无关. 后面要证明更精确的性质:如果在 $\boldsymbol{R}_0$ 下用 $\boldsymbol{A}=(a_{ij})$ 表示 $\boldsymbol{T}$,而在另一基底 $\boldsymbol{R}_1$ 下用 $\boldsymbol{B}=(b_{ij})$ 表示 $\boldsymbol{T}$,$\boldsymbol{R}_0$ 与 $\boldsymbol{R}_1$ 都是法正交的,则 $D_{\boldsymbol{A}}(s)\equiv D_{\boldsymbol{B}}(s)$.

3) 每一特征值都是实数. 事实上,从式(14) 推出

$$\boldsymbol{x}\cdot\boldsymbol{A}\boldsymbol{x}=s\boldsymbol{x}\cdot\boldsymbol{x}=s\|\boldsymbol{x}\|^2$$

故

$$s=\frac{\boldsymbol{x}\cdot\boldsymbol{A}\boldsymbol{x}}{\|\boldsymbol{x}\|^2}=\frac{\phi(\boldsymbol{x})}{\|\boldsymbol{x}\|^2}$$

若 $\boldsymbol{x}$ 是特征向量,则 $\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ 也是,所以特征值 s 是厄米形式 $\phi(\boldsymbol{x})$ 在属于 s 的单位特征向量 $\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ 处的值.

4) 属于两个不同特征值 s_1 与 s_2 的两个特征向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交.

事实上,从 $\boldsymbol{A}\boldsymbol{x}=s_1\boldsymbol{x}$ 与 $\boldsymbol{A}\boldsymbol{y}=s_2\boldsymbol{y}$,并考虑到极形式便推出: $\boldsymbol{y}\cdot\boldsymbol{A}\boldsymbol{x}=s_1\boldsymbol{y}\cdot\boldsymbol{x}$, $\boldsymbol{A}\boldsymbol{y}\cdot\boldsymbol{x}=s_2\boldsymbol{y}\cdot\boldsymbol{x}$. 故有

$$(s_1-s_2)\boldsymbol{y}\cdot\boldsymbol{x}=0,\text{或 } \boldsymbol{y}\cdot\boldsymbol{x}=0 \tag{16}$$

E_n 的情形　上述诸结果继续有效. 3) 的证明修改如下:命 s 是 $D_{\boldsymbol{A}}(s)=0$ 的根. 无论 s 是实数或复数,相应地存在一些数 x_i(可能为复数),不全为零,即是 n 个齐方程的解: $sx_i=\sum\limits_j a_{ij}x_j$,于是有

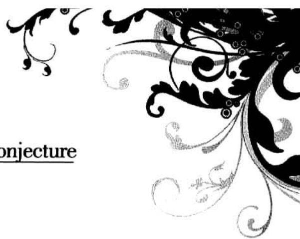
$$\sum_{i,j}\bar{x}_i a_{ij}x_j=s\sum_i\bar{x}_i x_i=s\sum_i|x_i|^2$$

由此得到

$$s=\frac{\sum\limits_{i,j}a_{ij}\bar{x}_i x_j}{\sum\limits_i|x_i|^2} \tag{17}$$

根据 $a_{ij}=a_{ji}$ 知,分子是实数,故 s 是实数,所以,对称矩阵的特征值是实数.





3 二次形式与厄米形式化为法式

算子 $\boldsymbol{y} = f(\boldsymbol{x})$ 叫酉算子,如果 $\| f(\boldsymbol{x}) \| = \| \boldsymbol{x} \|$.矩阵 $\boldsymbol{A} = (a_{ij})$ 叫酉矩阵,如果在法正交基底下由 $\boldsymbol{y} = \boldsymbol{Ax}$ 所定义的 C_n 的自同态是一个酉算子,即 $\| \boldsymbol{Ax} \| = \| \boldsymbol{x} \|$.

定理 2 矩阵 $\boldsymbol{A}$ 是酉的,必须且只需 $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I}$,$\boldsymbol{I}$ 是单位矩阵.

事实上,若 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I} \tag{18}$$

则由相伴矩阵 $\boldsymbol{A}^*$ 的定义有

$$\| \boldsymbol{Aa} \|^2 = \boldsymbol{Aa} \cdot \boldsymbol{Aa} = \boldsymbol{a} \cdot \boldsymbol{A}^*\boldsymbol{Aa} = \boldsymbol{a} \cdot \boldsymbol{a} = \| \boldsymbol{a} \|^2$$

反之,若对每一向量 $\boldsymbol{a}$,$\| \boldsymbol{Aa} \|^2 = \| \boldsymbol{a} \|^2$,则有

$$\| \boldsymbol{A}(\boldsymbol{a} + \boldsymbol{b}) \|^2 = (\boldsymbol{Aa} + \boldsymbol{Ab}) \cdot (\boldsymbol{Aa} + \boldsymbol{Ab}) = \| \boldsymbol{Aa} \|^2 + \| \boldsymbol{Ab} \|^2 + \boldsymbol{Aa} \cdot \boldsymbol{Ab} + \boldsymbol{Ab} \cdot \boldsymbol{Aa}$$

$$\| \boldsymbol{A}(\boldsymbol{a} + \boldsymbol{b}) \|^2 = \| \boldsymbol{a} + \boldsymbol{b} \|^2 = \| \boldsymbol{a} \|^2 + \| \boldsymbol{b} \|^2 + \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{a}$$

由此

$$\boldsymbol{Aa} \cdot \boldsymbol{Ab} + \boldsymbol{Ab} \cdot \boldsymbol{Aa} = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{a} \tag{19}$$

用 $\boldsymbol{b}' = i\boldsymbol{b}$ 代替(19)的 $\boldsymbol{b}$ 得

$$\boldsymbol{Aa} \cdot \boldsymbol{Ab} - \boldsymbol{Ab} \cdot \boldsymbol{Aa} = \boldsymbol{a} \cdot \boldsymbol{b} - \boldsymbol{b} \cdot \boldsymbol{a}$$

将上两等式相加得

$$\boldsymbol{Aa} \cdot \boldsymbol{Ab} = \boldsymbol{a} \cdot \boldsymbol{b} \tag{20}$$

因此,保范蕴涵保纯量积.于是得

$$\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{Aa} \cdot \boldsymbol{Ab} = \boldsymbol{a} \cdot \boldsymbol{A}^*\boldsymbol{Ab}$$

即是对任意 $\boldsymbol{a}$,$\boldsymbol{a} \cdot [\boldsymbol{A}^*\boldsymbol{Ab} - \boldsymbol{b}] = 0$.若将方括弧中的向量取作 $\boldsymbol{a}$,则对每一 $\boldsymbol{b}$ 有 $\boldsymbol{A}^*\boldsymbol{Ab} - \boldsymbol{b} = \boldsymbol{0}$,故得式(18).

所以酉矩阵是正则的,即其行列式 δ 非零,因为

$$\delta\bar{\delta} = 1, \ |\delta| = 1$$

E_n 的情形 若将 E_n 视为从点 O 引出的向量集合,则酉自同态是绕 O 的旋转.一个矩阵 $\boldsymbol{A}$ 如果关于法正交基底定义一个酉自同态,就叫做正交矩阵,这时,定理 2(证明相同)说:矩阵 $\boldsymbol{A}$ 是正交矩阵的充要条件是:$\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{I}$ 或 $\boldsymbol{A}' = \boldsymbol{A}^{-1}$.

证明变简单了:从式(19)出发,由于实纯量积可换,推出 $\boldsymbol{Aa} \cdot \boldsymbol{Ab} = \boldsymbol{a} \cdot \boldsymbol{b}$,即式(20),其余同上.

注

1) C_n 的酉变换构成一个群,叫做酉群;同样,E_n 的正交变换构成正交群.

这些群是一种特殊的群:每个群运算都由 n^2 个参数(即矩阵 $\boldsymbol{A}$ 的元素 a_{ij})所确定,对 C_n 这 n^2 个数是复数,对 E_n 是实数,对于运算 $O = O_1O_2$ 即群运算 O_1 和 O_2 之积,相应的参数可以解析地表为 O_1 与 O_2 的参数的函数(只需计算乘积矩阵的项). 这样的群叫做李群.

2) 酉矩阵的 n^2 个元素 a_{ij} 之间有某些关系,可以从式(18)得到:令 δ_{ij} 表示单位矩阵 $\boldsymbol{I}$ 的元素($\delta_{ii} = 1$;若 $i \neq j$,则 $\delta_{ij} = 0$),将式(18)用元素表出便得

$$\sum_s \bar{a}_{si}a_{sj} = \delta_{ij} \tag{21}$$

反之,式(21)蕴涵 $\boldsymbol{BA} = \boldsymbol{I}$,而 $\boldsymbol{B} = \boldsymbol{A}^*$,由此

$$\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{I} = \boldsymbol{AA}^* \tag{22}$$

(22)的第二个等式等价于

$$\sum_s a_{is}\bar{a}_{js} = \delta_{ij} \tag{23}$$

因此,方程组(21)与(23)等价. 在正交群的情形(a_{ij} 为实数),对于 $n = 3$,我们重新得到 9 个数构成一个正交矩阵的经典条件.

3) 给了两个法正交基底 $\boldsymbol{R}_0$ 与 $\boldsymbol{R}_1$,分别由向量 $\boldsymbol{e}_i$ 与 $\boldsymbol{e}'_i$ 组成,则存在一酉矩阵 $\boldsymbol{S}$,使得

$$\boldsymbol{e}'_i = \boldsymbol{S}\boldsymbol{e}_i$$

事实上,由式(6),若 $\boldsymbol{S} = (s_{ij})$ 使得

$$\boldsymbol{e}'_i = \boldsymbol{S}\boldsymbol{e}_i = \sum_j s_{ji}\boldsymbol{e}_j$$

由此推出

$$\boldsymbol{e}'_i \cdot \boldsymbol{e}'_k = \boldsymbol{S}\boldsymbol{e}_i \cdot \boldsymbol{S}\boldsymbol{e}_k = \sum_j \bar{s}_{ji}s_{jk} = \delta_{ik}$$

据(21),这表明 $\boldsymbol{S}$ 是酉矩阵.

如果把任何一个单位向量取作 $\boldsymbol{e}'_1$,则可确定一个法正交基(定理 1),从而可以确定一酉矩阵 $\boldsymbol{S}$,使得 $\boldsymbol{S}\boldsymbol{e}_1 = \boldsymbol{a}_1$ 是事先给出的单位向量.

这个结论(把酉矩阵换成正交矩阵)在 E_n 中仍有效,并且有几何解释.

定理 3 若 $\boldsymbol{A}$ 是厄米矩阵,$\boldsymbol{S}$ 是酉矩阵,则 $\boldsymbol{B} = \boldsymbol{SAS}^{-1}$ 是厄米矩阵.

关于法正交基 $\boldsymbol{R}_0$,视 $\boldsymbol{y} = \boldsymbol{Ax}$ 为厄米算子,我们来证明 $\boldsymbol{y} = \boldsymbol{Bx}$ 也是厄米算子,即要证

$$\boldsymbol{Bx} \cdot \boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{By}$$

或者

$$\boldsymbol{SAS}^{-1}\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{SAS}^{-1}\boldsymbol{y}$$

但由于有[①]

① 原文如觉不好理解,可改为:事实上由于(13)适用于 $\boldsymbol{A}$,而(20)适用于 $\boldsymbol{S}$ 和 $\boldsymbol{S}^{-1}$,所以 $\boldsymbol{SAS}^{-1}\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{S}^{-1}\boldsymbol{SAS}^{-1}\boldsymbol{x} \cdot \boldsymbol{S}^{-1}\boldsymbol{y} = \boldsymbol{AS}^{-1}\boldsymbol{x} \cdot \boldsymbol{S}^{-1}\boldsymbol{y} = \boldsymbol{S}^{-1}\boldsymbol{x} \cdot \boldsymbol{AS}^{-1}\boldsymbol{y} = \boldsymbol{SS}^{-1}\boldsymbol{x} \cdot \boldsymbol{SAS}^{-1}\boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{SAS}^{-1}\boldsymbol{y}$.



$$Au \cdot v = u \cdot Av$$

若令 $u = S^{-1}x, v = S^{-1}y$,则得

$$AS^{-1}x \cdot S^{-1}y = S^{-1}x \cdot AS^{-1}y$$

由于式(20)适用于酉矩阵 S 与 S^{-1},又有

$$SAS^{-1}x \cdot y = x \cdot SAS^{-1}y$$

同样,若 A 是对称矩阵,S 是正交矩阵,则 $B = SAS^{-1}$ 是对称矩阵.

注 若算子(15)关于法正交基底 R_0 由厄米矩阵 A 表示,关于基底 R_1 由矩阵 B 表示,则 $D_A(s) \equiv D_B(s)$.

事实上,设 S 是酉矩阵,它将 R_0 变成 R_1,使得 $B = (S^*)^{-1} \cdot AS^{-1} = SAS^{-1}$,由此,$D_B = [B - sI]$ 的行列式 $= [SAS^{-1} - SI]$ 的行列式 $= S[A - SI]S^{-1}$ 的行列式 $= D_A$.

厄米形式化为法式 设 $\phi = \sum_{i,j} a_{ij}\bar{x}_i x_j$ 是我们要处理的形式,$\phi(x) = x \cdot Ax, A = (a_{ij})$ 是厄米算子 $y = Ax = Tx$ 在法正交基底 R_0 下的表示矩阵,x_i 是 x 在 R_0 下的坐标,于是 $\phi(x)$ 是 R_0 下附属于 T 的厄米形式.

定理 A 在 C_n 内至少存在一法正交基底 R,使得若 ξ_i 是 x 在 R 下的坐标,则 $\phi(x) = x \cdot Tx$ 有如下表达式

$$\phi = \sum_{i=1}^{p} s_i \bar{\xi}_i \xi_i \quad p \leqslant n \tag{24}$$

诸 s_i 是实数.

令 $x_i = \sum_j s_{ij}\xi_j$ 矩阵 $S = (s_{ij})$ 是酉矩阵,则 $\phi = \sum_{s,t} b_{st}\bar{\xi}_s \xi_t$,这里 $B = (b_{st}) = S^*AS = S^{-1}AS$.

于是定理 A 的结论等价于:

定理 B 对于厄米矩阵 A,至少存在一酉矩阵 S,使得 $B = S^{-1}AS$ 有对角形式

$$B = \begin{pmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & s_n \end{pmatrix}$$

诸 s_i 是实数.

我们用下述引理来证明定理 A:

引理 若 a_1 是厄米算子 T 的一个单位特征向量,令 $x = x_1 a_1 + u$ 且 $a_1 \cdot u = 0$,则 $y = Tx$ 可分解成如下形式

$$y = Tx = s_1 x_1 a_1 + v$$

s_1 是 $\boldsymbol{a}_1$ 所属的特征值;并且还有下列性质:

1) $\boldsymbol{v} = \boldsymbol{T}\boldsymbol{u}$;2) 与 $\boldsymbol{u}$ 一样,$\boldsymbol{v}$ 也在 $\boldsymbol{a}_1$ 的正交子空间 $C_{n-1}(\boldsymbol{a}_1)$ 内;3)$\phi = \boldsymbol{x}\cdot\boldsymbol{T}\boldsymbol{x} = s_1\bar{x}x_1 + \boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}$,这里 $\boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}$ 是空间 $C_{n-1}(\boldsymbol{a}_1)$ 中的厄米形式.

事实上,$\boldsymbol{T}(\boldsymbol{x}) = \boldsymbol{T}(x_1\boldsymbol{a}_1 + \boldsymbol{u}) = x_1\boldsymbol{T}(\boldsymbol{a}_1) + \boldsymbol{T}(\boldsymbol{u}) = s_1x_1\boldsymbol{a}_1 + \boldsymbol{T}(\boldsymbol{u})$.又,$\boldsymbol{a}_1\cdot\boldsymbol{T}(\boldsymbol{u}) = \boldsymbol{T}(\boldsymbol{a}_1)\cdot\boldsymbol{u} = s_1\boldsymbol{a}_1\cdot\boldsymbol{u} = 0$.从而

$$\boldsymbol{x}\cdot\boldsymbol{T}\boldsymbol{x} = (x_1\boldsymbol{a}_1 + \boldsymbol{u})\cdot(s_1x_1\boldsymbol{a}_1 + \boldsymbol{T}\boldsymbol{u}) = s_1\bar{x}_1x_1 + \boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}$$

推论 令 $\boldsymbol{R}_1 = (\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n)$ 是 C_n 的法正交基底,而 $\boldsymbol{a}_1$ 是 $\boldsymbol{T}\boldsymbol{x}$ 的属于 s_1 的特征向量,则厄米形式 $\phi(\boldsymbol{x}) = \boldsymbol{x}\cdot\boldsymbol{T}\boldsymbol{x}$ 可分解成

$$\phi = s_1\bar{x}_1x_1 + \phi_1 \tag{25}$$

这里 $\phi_1 = \boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}$ 是至多有 $n-1$ 个变量的厄米形式,相应于算子 $\boldsymbol{T}$ 在 $C_{n-1}(\boldsymbol{a}_1)$ 内的迹.

事实上,若对于 $\boldsymbol{R}_1$ 有 $\boldsymbol{x} = (x_1,\cdots,x_n)$,则 $\phi_1 = \boldsymbol{u}\cdot\boldsymbol{T}\boldsymbol{u}$,这里 $\boldsymbol{u}$ 与 $\boldsymbol{T}\boldsymbol{u}$ 均属于坐标为 $x_2,\cdots,x_n$ 的空间 $C_{n-1}(\boldsymbol{a}_1)$.

讨论 $\boldsymbol{a}_1$ 与 s_1 的存在已肯定(s_1 可能为零),所以来研究 ϕ_1.

a) 若 $\phi_1 \equiv 0$,则定理 A 成立,这里 $p \leqslant 1$

$$\boldsymbol{R} = (\boldsymbol{b}_1 = \boldsymbol{a}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)$$

$\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n$ 是由 $\boldsymbol{a}_1\cdot\boldsymbol{x} = 0$ 确定的子空间 $C_{n-1}(\boldsymbol{a}_1)$ 中的任何法正交基底.对于 $k \geqslant 2$ 有 $\boldsymbol{T}\boldsymbol{b}_k = 0$;向量 $\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n$ 是 $\boldsymbol{T}$ 的属于特征值 $s = 0$ 的 $n-1$ 个特征向量.在这个基底 $\boldsymbol{R}$ 下,我们有 $\phi(\boldsymbol{x}) = \boldsymbol{x}\cdot\boldsymbol{T}\boldsymbol{x} = s_1\bar{\xi}_1\xi_1, x_1 = \xi_1, D(s) = (-1)^{n-1}\cdot(s_1 - s)s^{n-1}$.

b) 若 $\phi_1 \not\equiv 0$,选单位向量 $\boldsymbol{b}_2$ 作 $\boldsymbol{v} = \boldsymbol{T}\boldsymbol{u}$ 的特征向量,$\boldsymbol{T}$ 是 $C_{n-1}(\boldsymbol{a}_1)$ 中的厄米算子;$\boldsymbol{b}_2$ 是 $\boldsymbol{T}$ 在 C_n 中的特征向量,与 $\boldsymbol{b}_1 = \boldsymbol{a}_1$ 正交,所以关于基底 $\boldsymbol{R}_2(\boldsymbol{b}_2 = \boldsymbol{a}_1,\boldsymbol{b}_2 = \boldsymbol{a}_2,\boldsymbol{b}_3,\cdots,\boldsymbol{b}_n)$ 有

$$\phi = s_1\bar{x}_1x_1 + s_2\bar{x}_2x_2 + \phi_2$$

ϕ_2 是至多有 $n-2$ 个变量的厄米形式:$\phi_2 = \boldsymbol{\omega}\cdot\boldsymbol{T}\boldsymbol{\omega}$,$\boldsymbol{\omega}$ 属于由 $\boldsymbol{b}_1\cdot\boldsymbol{x} = 0$ 与 $\boldsymbol{b}_2\cdot\boldsymbol{x} = 0$ 确定的子空间 $C_{n-1}(\boldsymbol{b}_1,\boldsymbol{b}_2)$.

若 $\phi_2 \equiv 0$,则定理 A 成立,这时 $p \leqslant 2$,$\boldsymbol{R} = (\boldsymbol{a}_1 = \boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)$,$\boldsymbol{b}_3,\cdots,\boldsymbol{b}_n$ 是 $C_{n-2}(\boldsymbol{b}_1,\boldsymbol{b}_2)$ 的任何法正交基底.若 $\phi_2 \not\equiv 0$,继续利用前面的引理与推论,经过 $p \leqslant n$ 次后,便得法式(24).

注 指出下列几种不同的提法:

1) 利用定理 3 可以就矩阵进行讨论;容易获得定理 B 的形式;$\boldsymbol{a}_1$ 总是 $\boldsymbol{y} = \boldsymbol{T}\boldsymbol{x}$ 的单位特征向量,$\boldsymbol{R}_1$ 是法正交基底$(\boldsymbol{b}_1 = \boldsymbol{a}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n)$;$\boldsymbol{S}$ 是酉变换,它将 $\boldsymbol{R}_0$ 变成 $\boldsymbol{R}_1$,我们有

$$\phi(\boldsymbol{x}) = \sum_{i,j} a_{ij}\bar{x}_ix_j = \boldsymbol{x}\cdot\boldsymbol{A}\boldsymbol{x}(\text{关于 } \boldsymbol{R}_0),\boldsymbol{A} = (a_{ij})$$



$$\phi(\boldsymbol{x}) = \sum_{i,j} b_{ij}\bar{x}'_i x'_j = \boldsymbol{x} \cdot \boldsymbol{Bx}(\text{关于 } \boldsymbol{R}_1), \boldsymbol{B} = (b_{ij})$$

$$\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{AS}$$

由定理3知,$\boldsymbol{B}$ 是厄米矩阵,故 $\boldsymbol{Ba}_1 = s_1\boldsymbol{a}_1$ 蕴涵 $b_{11} = s_1, b_{1i} = b_{j1} = 0$ $(j \geqslant 2)$,$b_{11} = \bar{b}_{11}$.这就直接确定特征值 s_1 是实数;$\boldsymbol{B}$ 由一个$(n-1)\times(n-1)$型厄米矩阵 $\boldsymbol{B}' = (b'_{ij}), b'_{ij} = b_{ij}(2 \leqslant i \leqslant n, 2 \leqslant j \leqslant n)$ 在其左上角补以 s_1,其余补以0所构成.假定定理B的结论对 $n-1$ 为真,则存在$(n-1)\times(n-1)$型酉矩阵 $\boldsymbol{\theta}$,使得 $\boldsymbol{\theta}^{-1}\boldsymbol{B}'\boldsymbol{\theta}$ 是一个对角矩阵,对角线元素为实数 $s_2,\cdots,s_n$;在 $\boldsymbol{\theta}$ 的对角线左上角处补以1,其余补以0得到的 $n \times n$ 型矩阵 $\boldsymbol{T}$ 是 C_n 中的酉矩阵;同样 $\boldsymbol{S}_1 = \boldsymbol{TS}$ 也是酉矩阵;于是 $\boldsymbol{B}'_1 = \boldsymbol{S}_1^{-1}\boldsymbol{AS}$ 为对角厄米矩阵,这就证明了定理B.令 $s_1 = (s_{ij}^{(1)})$,代换 $x_i = \sum_1^n s_{ij}^{(1)}\xi_j$ 给出 $\phi = \sum_1^n s_i\bar{\xi}_i\xi_i$,改变脚标的顺序(这相当于作一酉变换),可以先写出非零项,从而得到定理A的形式.

2) 设 $\boldsymbol{\beta}_k$ 是上面定理中得到的基底$\boldsymbol{R}$ 的向量,它们是 $\boldsymbol{T}$ 的特征向量:$\boldsymbol{T\beta}_k = s_k\boldsymbol{\beta}_k$.假设它们的次序满足 $s_1 \geqslant s_2 \geqslant \cdots \geqslant s_p$,在定理 A 中若$p < n$,则 $s_{p+1},\cdots,s_n$ 是零.由(24)得

$$\begin{cases} s_1 = \max\limits_{\|x\|=1} \phi(\boldsymbol{x}), \text{即 } s_1 = \max \dfrac{\phi(\boldsymbol{x})}{\|\boldsymbol{x}\|} \\ s_2 = \max\limits_{\|x\|=1} \phi(\boldsymbol{x}), \boldsymbol{x}\cdot\boldsymbol{\beta}_1 = 0 \\ s_k = \max\limits_{\|x\|=1} \phi(\boldsymbol{x}), \boldsymbol{x}\cdot\boldsymbol{\beta}_j = 0, (1 \leqslant j \leqslant k-1) \end{cases} \tag{26}$$

于是,特征向量 $\boldsymbol{\beta}_k$ 的计算法在于寻求使(26)达到最大的向量 $\boldsymbol{\beta}_k$,并求 $s_k = \phi(\boldsymbol{\beta}_k)$.

我们可以由此得到定理A的证明,只需利用下面的结论:定义在紧致空间上的多变量连续函数的确在此空间中达到它的上、下确界.

我们有下面的结论:

定理 设 $\boldsymbol{\beta}_1$ 是单位向量,使 $\phi(\boldsymbol{x})$ 在 $\|\boldsymbol{x}\| = 1$ 的条件下取得最大值,$\boldsymbol{\beta}_2$ 是单位向量,使 $\phi(\boldsymbol{x})$ 在条件 $\|\boldsymbol{x}\| = 1$ 与 $\boldsymbol{x}\cdot\boldsymbol{\beta}_1 = 0$ 下取得最大值,继此以推,$\boldsymbol{\beta}_k(1 \leqslant k \leqslant n)$ 是单位向量,使 $\phi(\boldsymbol{x})$ 在条件 $\|\boldsymbol{x}\| = 1$ 与 $\boldsymbol{x}\cdot\boldsymbol{\beta}_j = 0, 1 \leqslant j \leqslant k-1$ 下取得最大值,若在法正交基底 $R(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_n)$ 下表出 $\boldsymbol{x} = (\xi_j)$,便得法式(24),诸 s_k 由(26)给出.

事实上,设在基底 $\boldsymbol{R}$ 下 $\phi(\boldsymbol{x}) = \sum_{i,j} b_{ij}\bar{\xi}_i\xi_j$,依定理3,$(b_{ij})$ 是厄米矩阵.我们有 $\phi(\boldsymbol{\beta}_1) = s_1 = b_{11}$,并且对于 $\|\boldsymbol{x}\| = 1, \phi(\boldsymbol{x}) \leqslant s_1$.由此可见,对任意的 $\boldsymbol{x}$ 有

$$h(\boldsymbol{x}) = \phi(\boldsymbol{x}) - s_1\|\boldsymbol{x}\|^2 \leqslant 0$$

$h(\boldsymbol{x})$ 显然是厄米形式.在 $j \geqslant 3$ 时令 ξ_j 为零,则 $h(\boldsymbol{x}) \leqslant 0$ 写成

$$h(\boldsymbol{x}) = b_{12}\bar{\xi}_1\xi_2 + b_{21}\bar{\xi}_2\xi_1 + b_{22}\bar{\xi}_2\xi_2 \leqslant 0$$

这个不等式对任何 ξ_1,ξ_2 成立．令 $\xi_2 = \lambda\xi_1$，λ 为实数，那么 $|\xi_1|^2[(b_{12} + b_{12})\lambda + b_{22}\lambda^2] \leqslant 0$ 对任意实数 λ 成立，这要求 $b_{12} + b_{21} = 0, b_{22} \leqslant 0$．对于 $\xi_2 = i\lambda\xi_1$，λ 为实数，也得 $b_{12} - b_{21} = 0$．最后必有 $b_{12} = b_{21} = 0$，同样对于 $j > 1$ 有 $b_{1j} = b_{j1} = 0$．故在 $\boldsymbol{R}$ 下所讨论形式可以写成

$$\phi(\boldsymbol{x}) = s_1\bar{\xi}_1\xi_1 + \phi_1(\boldsymbol{x})$$

这里 $\phi_1(\boldsymbol{x})$ 是厄米形式，只含有变量 $\xi_2,\cdots,\xi_n$．$\xi_1 = 0$ 时，ϕ 化为 ϕ_1，将上述方法用于 ϕ_1，于是，依次确定向量 $\boldsymbol{\beta}_k$ 作为取得大值的元素，在这些向量组成的基底 $\boldsymbol{R}$ 下，ϕ 必取法式(24)．

3) 如果将可换乘法 $\bar{x}_i x_j$ 换成关于加法左、右分配的任何运算，则将 $\phi = \sum a_{ij}\bar{x}_i x_j$ 化成法式(24) 的计算无须修改仍然有效．因此，考虑外积($\bar{x}_i \wedge x_j = - x_j \wedge \bar{x}_i$)，我们得到结论：代换 $x_i = \sum\limits_j s_{ij}\xi_j$ 把厄米形式 ϕ 化成法式(24)，也把形式 $\Psi = \sum a_{ij}\bar{x}_i \wedge x_j$ 化为法式 $\Psi = \sum s_i\bar{\xi}_i \wedge \xi_i$．

E_n 的情形　这时由于 x_i 与 $\bar{x}_i$ 相同，除了由此引起的修正外，上述结论与证明可以照搬过来，所以得到

定理　A) 在 E_n 中至少存在一个法正交基底 $\boldsymbol{R}$，使得如在 $\boldsymbol{R}$ 下有 $\boldsymbol{x} = (\xi_i)$，则对称二次形式

$$F = \sum_{i,j} a_{ij}x_i x_j = \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x} \quad (a_{ij} = a_{ji})$$

可化成

$$F = \sum_{i=1}^{p} s_i\xi_i^2 \quad p \leqslant n$$

诸 s_i 是实数，$\boldsymbol{R}$ 的向量 $\boldsymbol{\beta}_k$ 是对称算子 $\boldsymbol{y} = A\boldsymbol{x}$ 的特征向量．变换 $x_i = \sum\limits_j s_{ij}\xi_j$ 由一正交矩阵 $\boldsymbol{S} = (s_{ij})$ 实现．

B) 对于对称矩阵 $\boldsymbol{A} = (a_{ij})$，存在一正交矩阵 $\boldsymbol{S} = (s_{ij})$，使得 $\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ 是对角矩阵．

定义　1) 二次形式 F(或厄米形式 ϕ) 是正定的，如果对于任何非零向量 $\boldsymbol{x} = (x_i)$，有 $F > 0$(或 $\phi > 0$)．

2) F 或 ϕ 是半正定的，如果 $F \geqslant 0$(或 $\phi \geqslant 0$)．

定理　要使 F(或 ϕ) 是正定的，充要条件是它有最大秩 $p = n$，并且所有特征值 $s_k > 0(1 \leqslant k \leqslant n)$．

上述结果将用来研究直角坐标下圆锥曲线与二次曲面方程的化简问题．它



们的分类取决于形式的指数与秩①. 下面是关于二次曲面的亏格的几点说明: s 的方程的根 s_k 是满足下述条件的参数值 s: 在齐次坐标 x_1, x_2, x_3, x_4 下由 $F(x_1, x_2, x_3, x_4) + s\sum_{i=1}^{3} x_i^2 = 0, x_4 = 0$ 所确定的(束中的)无限远圆锥曲线已经分解; 根 s_k 首先可表成如下形式: 与二次形式下相应的对称变换 $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ 的特征值 s_k, 是使得形式 $F_s = F(x_1, x_2, \cdots, x_n) + s\sum_{i=1}^{n} x_i^2$ 的秩小于 n 的那些值, 并且只能是那些值. 然后利用这个结果指出: 秩 $p \leqslant 2$ 的形式显然是两个实线性形式或共轭线性形式的乘积.

① 秩为 p, 指数为 $q(0 \leqslant q \leqslant p)$ 的二次形式的法式为 $F = -\xi_1^2 - \cdots - \xi_q^2 + \xi_{q+1}^2 + \cdots + \xi_p^2$.

第六讲　典型群①

1　体 K 上 n 元线性群

令 $K^n = E$ 是向量$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$组成的向量空间,$\boldsymbol{x}$ 的 n 个分量属于可换体 K.

考虑线性变换 $\boldsymbol{x} \to \boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$,定义为

$$\begin{cases} x'_1 = a_{11}x_1 + \cdots + a_{1n}x_n \\ x'_2 = a_{21}x_1 + \cdots + a_{2n}x_n \\ \quad\vdots \\ x'_n = a_{n1}x_1 + \cdots + a_{nn}x_n \end{cases}$$

诸系数 a_{ij} 属于体 K,并且相应的行列式不为零.所有这样的线性变换构成一个群,称为 K 上的 n 元线性群,记成 $GL(n,K)$.

特别,行列式为1的所有线性变换构成 $GL(n,K)$ 的子群,称为 n 元特殊线性群,或么模群,记成 $SL(n,K)$.

本文要讨论的群是一般线性群的所有子群,K 是复数体或实数体.按H·外尔的命名,这些群称为典型群.至于任意体的情形是近年研究的对象,其主要结果已汇集于范·德·瓦尔登的书《Die Gruppentheoretische Methode in der Quantenmechanik》与《Gruppen von Linearen Transformationen》和J·迪厄多内的书《Sur les Groupes Classiques》与《La géométrie des Groups Classiques》中.

2　对合线性变换

这种线性变换与它的逆重合,即它的表示矩阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}^2 = \boldsymbol{I},\text{其中 } \boldsymbol{I} \text{ 表示单位矩阵}$$

根据上面关系,$(\det \boldsymbol{A})^2 = 1$,故 $\boldsymbol{A}$ 是可逆的.

---

① 原作者 L. Lesieur(普瓦蒂埃理学院教授).



这些矩阵关于矩阵的加法与乘法具有环的结构,所以关系 $\boldsymbol{A}^2=\boldsymbol{I}$ 可写成

$$\boldsymbol{A}^2-\boldsymbol{I}=\boldsymbol{O},\text{即}(\boldsymbol{A}-\boldsymbol{I})(\boldsymbol{A}+\boldsymbol{I})=\boldsymbol{O}$$

这就使我们来考虑满足下述条件的向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$

$$(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0},\text{即 }\boldsymbol{Ax}=\boldsymbol{x}$$

$$(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{y}=\boldsymbol{0},\text{即 }\boldsymbol{Ay}=-\boldsymbol{y}$$

满足前一关系的向量构成 E 的一个向量子空间 V,后者构成一个向量子空间 W.这两个子空间 V 与 W 没有公共的非零向量,因为:

$\boldsymbol{x}=\boldsymbol{y}$ 蕴涵 $\boldsymbol{Ax}=\boldsymbol{Ay}$,即 $\boldsymbol{x}=-\boldsymbol{y}$,由此 $2\boldsymbol{x}=\boldsymbol{0}$,即 $\boldsymbol{x}=\boldsymbol{0}$(对于示性数不是 2 的体,可以施行除以 2 的运算).

另一方面,V 与 W 生成空间 E,即 E 的每一向量 $\boldsymbol{v}$ 可写成

$$\boldsymbol{v}=\boldsymbol{x}+\boldsymbol{y}\quad \boldsymbol{x}\in V,\boldsymbol{y}\in W$$

事实上,我们有

$$\boldsymbol{v}=\frac{1}{2}(\boldsymbol{v}+\boldsymbol{Av})+\frac{1}{2}(\boldsymbol{v}-\boldsymbol{Av})$$

这里 $\boldsymbol{A}(\boldsymbol{v}+\boldsymbol{Av})=\boldsymbol{Av}+\boldsymbol{v}$,故 $\boldsymbol{v}+\boldsymbol{Av}\in V$,而 $\boldsymbol{A}(\boldsymbol{v}-\boldsymbol{Av})=\boldsymbol{Av}-\boldsymbol{v}=-(\boldsymbol{v}-\boldsymbol{Av})$,故 $\boldsymbol{v}-\boldsymbol{Av}$ 属于 W.

这两条性质可表成如下形式

$$V\cap W=0,V\cup W=E\text{①}$$

并说 V 与 W 互补.

因此有下面定理:

定理 1 设 A 是对合线性变换,则存在两个互补的子空间 V 与 W,使得对 V 的每个向量 $\boldsymbol{x}$ 有 $\boldsymbol{Ax}=\boldsymbol{x}$,对 W 的每个向量 $\boldsymbol{x}$ 有 $\boldsymbol{Ax}=-\boldsymbol{x}$.

若 V 是 p 维,则 W 是 $n-p$ 维;在 V 中选基底 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_p\}$,$W$ 中选 $\{\boldsymbol{e}_{p+1},\cdots,\boldsymbol{e}_n\}$,则变换方程取如下形式

$$\begin{cases}x'_1=x_1\\x'_2=x_2\\\vdots\\x'_p=x_p\\x'_{p+1}=-x_{p+1}\\\vdots\\x'_n=-x_n\end{cases}$$

E^3 中的例子.

① 记号 $\cup$ 不是集论中并集的记号,而是格伦中和的记号(参看 L. Dubreil Jacotin, L. Lesieur, R. Croisot,《Théorie des treillis》,1953,p.9 与 25,例 6).

$$\begin{cases} x'_1 = x_1 \\ x'_2 = x_2 \\ x'_3 = -x_3 \end{cases}$$（关于平面$(\boldsymbol{e}_1,\boldsymbol{e}_2)$平行于$\boldsymbol{e}_3$的对称变换）.

3 正交群

一个线性变换如果使二次形式

$$\varphi(\boldsymbol{x}) = x_1^2 + x_2^2 + \cdots + x_n^2$$

保持不变,就称为正交变换.

相应的矩阵 $\boldsymbol{A}$ 叫正交矩阵.一个矩阵 $\boldsymbol{A}$ 是正交的,必须且只需 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{I}$$

这里 $\boldsymbol{A}^{\mathrm{T}}$ 表示 $\boldsymbol{A}$ 的转置矩阵(由 $\boldsymbol{A}$ 关于主对角线对称而得的矩阵).

对 $n = 2$ 予以验证.要使变换

$$x'_1 = a_{11}x_1 + a_{12}x_2$$

$$x'_2 = a_{21}x_1 + a_{22}x_2$$

对于任意的 x_1 与 x_2 满足 $x'^2_1 + x'^2_2 \equiv x_1^2 + x_2^2$,条件是

$$\begin{cases} a_{11}^2 + a_{21}^2 = 1, a_{12}^2 + a_{22}^2 = 1 \\ a_{11}a_{12} + a_{21}a_{22} = 0 \end{cases}$$

这条件可表成

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\text{是单位矩阵}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

由此推出

$$\det \boldsymbol{A} = \pm 1$$

所以,正交矩阵是可逆的.此外,它的逆是其转置 $\boldsymbol{A}^{\mathrm{T}}$,并满足:$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{I}$,故 $\boldsymbol{A}^{\mathrm{T}}$ 也是正交矩阵.

因为两个正交变换的乘积仍是正交变换,故所有正交变换构成一个群,称为正交群,记为 $O(n,K)$.有两种正交变换:行列式为 $+1$ 的称为旋转,构成 $O(n,K)$ 的子群 $O^+(n,K)$,而行列式为 -1 的则称为翻转(当然不构成子群).

若 K 是复数体,则实数体 $\mathbf{R}$ 上的实正交群 $O(n,\mathbf{R})$ 显然是 $O(n,K)$ 的子群.

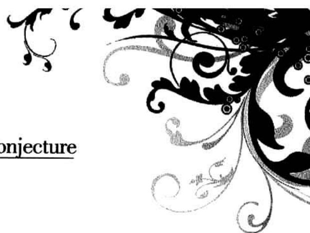
性质 正交变换保持两个向量的纯量积.

向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的纯量积表成

$$f(\boldsymbol{x},\boldsymbol{y}) = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

它是向量的坐标的对称双线性形式.





事实上,上面的形式可写成矩阵的乘积

$$f(\boldsymbol{x},\boldsymbol{y}) = (x_1 \quad x_2 \quad \cdots \quad x_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$$

经过变换 $\boldsymbol{A}$,它变成(根据结合律)

$$(\boldsymbol{AX})^{\mathrm{T}}\boldsymbol{AY} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}})(\boldsymbol{AY}) = \boldsymbol{X}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}$$

子空间 V 的共轭子空间 V^* 是满足下述条件的向量 $\boldsymbol{y}$ 所成的子空间

$$对每一\ \boldsymbol{x} \in V, f(\boldsymbol{x},\boldsymbol{y}) = 0$$

V^* 的确是子空间,因为 f 双线性;而上述关系是对称的,故 V 也是 V^* 的共轭子空间.

若 V 只由一个向量 $\boldsymbol{x}$ 生成,则 V^* 是正交超平面.更一般,若 V 是 p 维,则 V^* 是 $n-p$ 维(事实上,V 由 p 个无关向量 $\boldsymbol{e}_1,\cdots,\boldsymbol{e}_p$ 确定,V^* 是 p 个无关超平面 $f(\boldsymbol{e}_i,\boldsymbol{y}) = 0$ 的交,$i = 1,2,\cdots,p$).

当 K 是实数体时,我们有 $V \cap V^* = 0$,故两个共轭的子空间是互补的.当 K 是复数体时,结论不真,这时,存在子空间 V,使得 $V \cap V^* \neq 0$,称为迷向子空间.

4 正交对合变换

设 $\boldsymbol{A}$ 是正交对合矩阵(故 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2 = \boldsymbol{I}$,由此 $\boldsymbol{A} = \boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}}$).这是一种特殊的对合变换,定理1适用.我们来证明该定理中的 V 和 W 是共轭子空间,从而使结论更加确切.

$\boldsymbol{x} \in V$ 与 $\boldsymbol{y} \in W$ 的纯量积是

$$f(\boldsymbol{x},\boldsymbol{y}) = f(A\boldsymbol{x},A\boldsymbol{y}) = f(\boldsymbol{x},-\boldsymbol{y}) = -f(\boldsymbol{x},\boldsymbol{y})$$

由此 $2f(\boldsymbol{x},\boldsymbol{y}) = 0, f(\boldsymbol{x},\boldsymbol{y}) = 0$,只要体 K 的示性数不是2.

由此推出 $y \in V^*$,故 $W \subseteq V^*$.但 W 与 V^* 有同样的维数 $n-p$,所以 $W = V^*$.因此,V 与 W 共轭.

此外,我们知道 $V \cap W = 0$,故 $V \cap V^* = 0$,所以 V 不是迷向子空间.W 也一样.由此有:

定理2 若 $\boldsymbol{A}$ 是对合正交变换,则存在非迷向子空间 V(可能只有零元),使得对于 V 的每个向量 $\boldsymbol{x}$ 有 $\boldsymbol{Ax} = \boldsymbol{x}$,对于共轭子空间 V^* 的每个向量 $\boldsymbol{x}$ 有 $\boldsymbol{Ax} = -\boldsymbol{x}$.

在 V 与 W 内各取一个法正交基底,便得 E 的一个法正交基底,使得变换方程取式(1)的形式,该变换是旋转或翻转依 $n-p$ 是偶数或奇数而定.

5　对称变换

特别,当 V 是 $n-1$ 维时,对合正交变换使 V 保持不变,称为关于超平面 V 的对称变换(已知 V 不是迷向子空间).初等几何中由对称变换生成正交群的著名定理,在复数体 K 或实数体 $\mathbf{R}$ 的情形,已由 E · 嘉当推广至任意的 n,J · 迪厄多内又推广至示性数不是 2 的任意体.即下述定理成立.

定理 3　任意 n 元正交变换是至多 n 个对称变换的积.

对 n 用归纳法就实数体证明.$n=1$ 时显然真,设 $n-1$ 时定理真.若存在一向量 $\boldsymbol{x}\neq\mathbf{0}$,使得 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}$,则 $\boldsymbol{A}$ 使得与 $\boldsymbol{x}$ 共轭的整个超平面 H 不变,$\boldsymbol{A}$ 在 H 上的限制是正交变换,它是至多 $n-1$ 个对称变换的积.这就在 E 中定义了一些对称变换,至多 $n-1$ 个.若没有一个非零向量是不变的,考虑任一向量 $\boldsymbol{x}_0$,则 $\boldsymbol{A}\boldsymbol{x}_0\neq\boldsymbol{x}_0$,$\boldsymbol{x}_0$ 满足

$$\varphi(\boldsymbol{x}_0)=f(\boldsymbol{x}_0,\boldsymbol{x}_0)=\varphi(\boldsymbol{A}\boldsymbol{x}_0)$$

方程为 $f(\boldsymbol{x},\boldsymbol{x}_0-\boldsymbol{A}\boldsymbol{x}_0)=0$ 的超平面定义一个对称变换 S,把 $\boldsymbol{x}_0$ 变成 $\boldsymbol{A}\boldsymbol{x}_0\neq 0$.积变换 $\boldsymbol{SA}$ 使 $\boldsymbol{x}_0$ 不变,故是至多 $n-1$ 个对称变换的积,所以 $\boldsymbol{A}$ 是至多 n 个对称变换的积.

6　酉　群

任取空间 $E=K^n$,这里 K 是复数体,现在考虑厄米形

$$\varphi(\boldsymbol{x})=x_1\overline{x}_1+x_2\overline{x}_2+\cdots+x_n\overline{x}_n$$

其中,$\overline{x}_i$ 表示 x_i 的共轭复数.酉变换是保持 $\varphi(\boldsymbol{x})$ 不变的线性变换,它的矩阵 $\boldsymbol{A}$ 满足关系

$$\overline{\boldsymbol{A}}^{\mathrm{T}}\cdot\boldsymbol{A}=\boldsymbol{I}$$

这里 $\overline{\boldsymbol{A}}^{\mathrm{T}}$ 表示 $\boldsymbol{A}$ 的共轭转置矩阵.故矩阵 $\boldsymbol{A}$ 可逆,其逆是 $\overline{\boldsymbol{A}}^{\mathrm{T}}$,满足 $\boldsymbol{A}\overline{\boldsymbol{A}}^{\mathrm{T}}=\boldsymbol{I}$,故也是厄米矩阵.两个酉变换的乘积仍是酉变换,故所有的酉变换构成一个群:酉群 $U(n,K)$.

一个酉变换使下列表达式不变

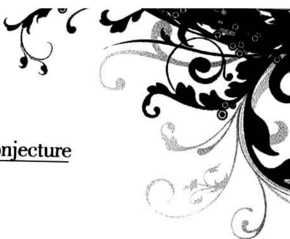
$$f(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_i\overline{y}_i$$

上式称为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的厄米纯量积,它是这两个向量的双线性形式,非对称,但满足

$$f(\boldsymbol{x},\boldsymbol{y})=\overline{f(\boldsymbol{y},\boldsymbol{x})}$$

故条件 $f(\boldsymbol{x},\boldsymbol{y})=0$ 与 $f(\boldsymbol{y},\boldsymbol{x})=0$ 等价.这些性质使我们能够定义共轭子空间





V 与 V^*,并将定理 3 扩张至酉群,得出任意酉变换的构造.

行列式等于 1 的酉变换构成酉群与特殊线性群的子群:特殊酉群 $S\bigcup(n, K)$.

7 辛群

假定空间 E 是偶维:$n = 2m$.向量 $\boldsymbol{x}$ 的坐标是:$x_1, x'_1, x_2, x'_2, \cdots, x_m, x'_m$.

对于两个向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$,我们考虑相应的

$$f(\boldsymbol{x}, \boldsymbol{y}) = x_1 y'_1 - x'_1 y_1 + x_2 y'_2 - x'_2 y_2 + \cdots + x_m y'_m - x'_m y_m$$

这是两个向量的交错双线性形式

$$f(\boldsymbol{y}, \boldsymbol{x}) = -f(\boldsymbol{x}, \boldsymbol{y})$$

保持 $f(\boldsymbol{x}, \boldsymbol{y})$ 不变的线性变换称为辛变换,因此它由下式定义

$$f(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{A}\boldsymbol{y})$$

相应的矩阵 $\boldsymbol{A}$ 称为辛矩阵.两个辛变换的积仍是辛变换.为证明辛变换构成一个群,我们来证明辛变换是可逆的.

$n = 2$ 时,辛变换就是保持定向面积不变的变换,所以是特殊线性群 $SL(2, K)$ 中的变换.在 $n = 2m$ 的一般情形,考虑矩阵 $\boldsymbol{A}$ 的列向量,即

$$\boldsymbol{\alpha}_1 = \boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{\alpha}'_1 = \boldsymbol{A}\boldsymbol{e}'_1, \cdots, \boldsymbol{\alpha}_m = \boldsymbol{A}\boldsymbol{e}_m, \boldsymbol{\alpha}'_m = \boldsymbol{A}\boldsymbol{e}'_m$$

从关系 $f(\boldsymbol{e}_i, \boldsymbol{e}_j) = 0, f(\boldsymbol{e}_i, \boldsymbol{e}'_j) = \delta_{ij}, f(\boldsymbol{e}'_i, \boldsymbol{e}'_j) = 0$(这里 δ_{ij} 当 $i = j$ 时为 1,$i \neq j$ 时为 0)推出

$$f(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j) = 0, f(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}'_j) = \delta_{ij}, f(\boldsymbol{\alpha}'_i, \boldsymbol{\alpha}'_j) = 0$$

这是 $\boldsymbol{A}$ 为辛矩阵的充要条件.而向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}'_1, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\alpha}'_m$ 是无关的,因为若存在若干 λ 使得 $\boldsymbol{x} = \lambda_1 \boldsymbol{\alpha}_1 + \lambda'_1 \boldsymbol{\alpha}'_1 + \cdots + \lambda_m \boldsymbol{\alpha}_m + \lambda'_m \boldsymbol{\alpha}'_m$ 为零,则 $f(\boldsymbol{x}, \boldsymbol{\alpha}'_1) = \lambda_1 = 0$,同样对 $1 \leqslant i \leqslant m, \lambda_i = \lambda'_i = 0$.故矩阵 $\boldsymbol{A}$ 可逆,而且它的逆也保持 $f(\boldsymbol{x}, \boldsymbol{y})$ 不变,所以也是辛矩阵.因此,辛变换构成一个群,称为辛群 $SP(n, K)$.

如同对正交群一样,我们还可定义共轭子空间 V 与 V^* 的概念,以及满足 $V \cap V^* \neq 0$ 的迷向子空间 V 的概念.

对合辛变换 利用同样的证法,定理 2 可推广至辛群.但我们可以补充说明一点:对合辛变换的结构中出现的空间 V 与 W 都是偶数维空间.

事实上,V 与 W 都不迷向,而奇数维子空间只能是迷向的:一个由基底 $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{2p+1}$ 确定的奇维空间 U,实际上含有非零向量

$$\boldsymbol{x} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_{2p+1} \boldsymbol{e}_{2p+1}$$

属于 U 的共轭子空间.因为一组 $2p + 1$ 个齐次方程

$$f(\boldsymbol{x}, \boldsymbol{e}_i) = \lambda_1 f(\boldsymbol{e}_1, \boldsymbol{e}_i) + \cdots + \lambda_{2p+1} f(\boldsymbol{e}_{2p+1}, \boldsymbol{e}_i) = 0$$

$$i = 1, 2, \cdots, 2p + 1$$

有奇数阶反对称行列式,故行列式为零,所以方程组有非零解 $\lambda_1,\lambda_2,\cdots,\lambda_{2p+1}$,从而给出非零向量.

特别,任何一维子空间以及任何超平面都是迷向子空间.

8 辛横截

前面的论证表明不存在辛"对称变换",但如同把正交对称变换视为保持超平面的向量不变的正交变换一样,我们也可寻求使得超平面的向量保持不变的辛变换.

设 A 是这样的变换,H 是超平面,$\boldsymbol{x}_0$ 是不在 H 上的向量;对于每个向量 $\boldsymbol{x}$ 相应地有 H 的向量 $\boldsymbol{h}$,使得 $\boldsymbol{x}=\boldsymbol{h}+r(\boldsymbol{x})\boldsymbol{x}_0$,由此 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{h}+r(\boldsymbol{x})\boldsymbol{A}\boldsymbol{x}_0$,相减得

$$\boldsymbol{A}\boldsymbol{x}-\boldsymbol{x}=r(\boldsymbol{x})\boldsymbol{a},\boldsymbol{a}=\boldsymbol{A}\boldsymbol{x}_0-\boldsymbol{x}_0$$

这个向量 $\boldsymbol{a}$ 是 H 的共轭向量,因为由 A 的定义,对 H 的任意向量 $\boldsymbol{e}$ 有

$$f(\boldsymbol{a},\boldsymbol{e})=f(\boldsymbol{A}\boldsymbol{x}_0-\boldsymbol{x}_0,\boldsymbol{e})=f(\boldsymbol{A}\boldsymbol{x}_0,\boldsymbol{e})-f(\boldsymbol{x}_0,\boldsymbol{e})=0$$

但由于任何向量都产生迷向子空间,所以 $\boldsymbol{a}$ 属于 H.因此,对于任何这样的变换 $\boldsymbol{A}$,存在 H 的共轭向量 $\boldsymbol{a}$ 与不属于 H 的向量 $\boldsymbol{x}_0$,使得 $\boldsymbol{x}=\boldsymbol{h}+r(\boldsymbol{x})\boldsymbol{x}_0,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}+r(\boldsymbol{x})\boldsymbol{a}$.这两个公式就确定了所求的变换.反之,由这两个公式定义的变换(其中 $\boldsymbol{x}_0$ 不在 H 上,$\boldsymbol{a}$ 是 H 的共轭向量),正好是一个辛变换,保持 H 的每一向量不变

$$\begin{aligned}f(\boldsymbol{A}\boldsymbol{x},\boldsymbol{A}\boldsymbol{y})=&f(\boldsymbol{x},\boldsymbol{y})+r(\boldsymbol{y})f(\boldsymbol{x},\boldsymbol{a})+\\&r(\boldsymbol{x})f(\boldsymbol{a},\boldsymbol{y})+r(\boldsymbol{x})r(\boldsymbol{y})f(\boldsymbol{a},\boldsymbol{x})=\\&f(\boldsymbol{x},\boldsymbol{y})\end{aligned}$$①

设 $2m-1$ 个向量 $\boldsymbol{e}_i$ 构成 H 的一个基底(取 $\boldsymbol{e}_1=\boldsymbol{a}$),向量 $\boldsymbol{e}_{2m}$ 在 H 之外(取 $\boldsymbol{e}_{2m}=\boldsymbol{x}_0$);上面的公式给出 $\boldsymbol{h}=x_1\boldsymbol{e}_1+\cdots+x_{2m-1}\boldsymbol{e}_{2m-1},r(\boldsymbol{x})=x_{2m}$,故 $\boldsymbol{A}\boldsymbol{x}$ 的坐标为

$$X_1=x_1+x_{2m},X_2=x_2,\cdots,X_{2m-1}=x_{2m-1}$$

$$X_{2m}=x_{2m}$$

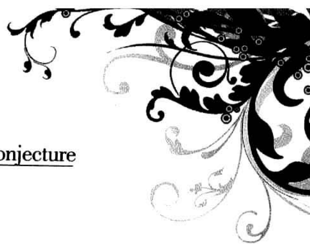
如上得到的变换称为超平面 H 的辛横截.上面的方程表明 $\boldsymbol{A}$ 的行列式为 1.

如果两个向量 $\boldsymbol{x}_0$ 和 $\boldsymbol{y}_0$ 满足 $f(\boldsymbol{x}_0,\boldsymbol{y}_0)\neq 0$,我们可以确定一个辛横截,把 $\boldsymbol{x}_0$ 变为 $\boldsymbol{y}_0$.事实上,按上面的结果,超平面 H 应是 $\boldsymbol{y}_0-\boldsymbol{x}_0$ 的共轭子空间,此超平面不含 $\boldsymbol{x}_0$,$\boldsymbol{A}$ 定义如下

---

① 注意,由于 $\boldsymbol{a}\in H$ 是 H 的共轭向量,所以 $f(\boldsymbol{a},\boldsymbol{a})=0$.此外,由于 $\boldsymbol{x}=\boldsymbol{h}+r(\boldsymbol{x})\boldsymbol{x}_0,\boldsymbol{y}=\boldsymbol{k}+r(\boldsymbol{y})\boldsymbol{x}_0,\boldsymbol{h},\boldsymbol{k}\in H$,所以 $r(\boldsymbol{y})f(\boldsymbol{x},\boldsymbol{a})=r(\boldsymbol{x})r(\boldsymbol{y})f(\boldsymbol{x}_0,\boldsymbol{a}),r(\boldsymbol{x})f(\boldsymbol{a},\boldsymbol{y})=r(\boldsymbol{x})r(\boldsymbol{y})f(\boldsymbol{a},\boldsymbol{x}_0)=-r(\boldsymbol{x})r(\boldsymbol{y})\cdot f(\boldsymbol{x}_0,\boldsymbol{a})$.





$$\boldsymbol{x}=\boldsymbol{h}+r(\boldsymbol{x})\boldsymbol{x}_0,\boldsymbol{A}\boldsymbol{x}=\boldsymbol{h}+r(\boldsymbol{x})\boldsymbol{y}_0$$

9 辛群的生成

定理 4 每个辛变换 $\boldsymbol{A}$ 是有限多个辛横截的积.

先假设对每个 $\boldsymbol{x}$,我们有 $f(\boldsymbol{x},\boldsymbol{A}\boldsymbol{x})=0$. 于是 $\boldsymbol{A}$ 是对合变换,因为从关系 $f(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{A}(\boldsymbol{x}+\boldsymbol{y}))=0$ 有 $f(\boldsymbol{x},\boldsymbol{A}\boldsymbol{y})=f(\boldsymbol{A}\boldsymbol{x},\boldsymbol{y})=f(\boldsymbol{A}\boldsymbol{x},\boldsymbol{A}^2\boldsymbol{y})$;由于向量 $\boldsymbol{A}\boldsymbol{x}$ 任意,故对任意 $\boldsymbol{y}$,$\boldsymbol{A}^2\boldsymbol{y}=\boldsymbol{y}$,即 $\boldsymbol{A}^2=\boldsymbol{I}$. 反之亦然,从 $f(\boldsymbol{x},\boldsymbol{y})=-f(\boldsymbol{y},\boldsymbol{x})$ 推出 $f(\boldsymbol{x},\boldsymbol{A}\boldsymbol{x})=0$. 因此有下列结论:

引理 1 如果对每个 $\boldsymbol{x}$ 有 $f(\boldsymbol{x},\boldsymbol{A}\boldsymbol{x})=0$,则 $\boldsymbol{A}$ 是对合变换;反之亦然.

在这种情形,可以考虑出现在对合辛变换的结构中的两个子空间 V 与 $W=V^*$. 设 $\boldsymbol{A}\neq\boldsymbol{I}$,则有 $W\neq 0$. 因 W 偶维,并且不是迷向子空间,所以 W 内存在两个向量 $\boldsymbol{x}_0$ 与 $\boldsymbol{y}_0$,使得 $f(\boldsymbol{x}_0,\boldsymbol{y}_0)\neq 0$. 从 $\boldsymbol{A}\boldsymbol{x}_0=-\boldsymbol{x}_0$ 这就表明 $\boldsymbol{x}\in W$ 推出 $f(\boldsymbol{A}\boldsymbol{x}_0,\boldsymbol{y}_0)\neq 0$,所以存在辛横截 T 将 $\boldsymbol{A}\boldsymbol{x}_0$ 变成 $\boldsymbol{y}_0$. 同样,$f(\boldsymbol{y}_0,\boldsymbol{x}_0)\neq 0$ 蕴涵存在辛横截 $\boldsymbol{T}'$ 将 $\boldsymbol{y}_0$ 变成 $\boldsymbol{x}_0$. 由此推出 $\boldsymbol{T}'\boldsymbol{T}\boldsymbol{A}$ 使 $\boldsymbol{x}_0$ 不变.

我们要证明,在一般情形,这结果仍真:

引理 2 设 $\boldsymbol{A}$ 是辛变换,则存在两个辛横截 $\boldsymbol{T}$ 与 $\boldsymbol{T}'$,使得 $\boldsymbol{T}'\boldsymbol{T}\boldsymbol{A}$ 至少保持一个向量 $\boldsymbol{x}_0$ 不变.

若 $\boldsymbol{A}$ 为对合变换,引理 2 就是刚才证明的性质. 若 $\boldsymbol{A}$ 不是对合变换,则由引理 1,至少存在一向量 $\boldsymbol{x}_0$,使得 $f(\boldsymbol{x}_0,\boldsymbol{A}\boldsymbol{x}_0)\neq 0$. 故存在辛横截 $\boldsymbol{T}$ 将 $\boldsymbol{A}\boldsymbol{x}_0$ 变成 $\boldsymbol{x}_0$,因此,积变换 $\boldsymbol{T}\boldsymbol{A}$ 使 $\boldsymbol{x}_0$ 不变.

现在设 $\boldsymbol{B}$ 是一辛变换,使 $\boldsymbol{x}_0$ 保持不变.

a) 先设存在 $\boldsymbol{y}_0$,使得

$$f(\boldsymbol{x}_0,\boldsymbol{y}_0)\neq 0,f(\boldsymbol{y}_0,\boldsymbol{B}\boldsymbol{y}_0)\neq 0$$

则存在辛横截 $\boldsymbol{T}_1$ 变 $\boldsymbol{B}\boldsymbol{y}_0$ 成 $\boldsymbol{y}_0$,积变换 $\boldsymbol{T}_1\boldsymbol{B}$ 使 $\boldsymbol{y}_0$ 不变. 向量 $\boldsymbol{x}_0$ 对这个积变换也是不变的. 为证明这点,只需验明它在辛横截 T_1 的不变超平面上,即它是 $\boldsymbol{B}\boldsymbol{y}_0-\boldsymbol{x}_0$ 的共轭向量. 这由下面等式推出

$$f(\boldsymbol{x}_0,\boldsymbol{B}\boldsymbol{y}_0-\boldsymbol{y}_0)=f(\boldsymbol{x}_0,\boldsymbol{B}\boldsymbol{y}_0)-f(\boldsymbol{x}_0,\boldsymbol{y}_0)=$$
$$f(\boldsymbol{B}\boldsymbol{x}_0,\boldsymbol{B}\boldsymbol{y}_0)-f(\boldsymbol{x}_0,\boldsymbol{y}_0)=0$$

我们刚才不仅就所考虑的情形证明了引理 2,而且得到更完善的结论:

引理 3 若存在向量 $\boldsymbol{y}_0$ 使得 $f(\boldsymbol{x}_0,\boldsymbol{y}_0)\neq 0,f(\boldsymbol{y}_0,\boldsymbol{B}\boldsymbol{y}_0)\neq 0$,则存在辛横截的积变换,再乘以 $\boldsymbol{A}$ 就使两个向量 $\boldsymbol{x}_0,\boldsymbol{y}_0$ 不变.

b) 现在设,对于每个满足 $f(\boldsymbol{x}_0,\boldsymbol{y}_0)\neq 0$ 的向量 $\boldsymbol{y}_0$,有 $f(\boldsymbol{y}_0,\boldsymbol{B}\boldsymbol{y}_0)=0$. 我们要利用一个辛横截,它将 $\boldsymbol{B}\boldsymbol{y}_0$ 不再变成 $\boldsymbol{y}_0$,而变成 $\boldsymbol{x}_0+\boldsymbol{y}_0$. 这个变换 $\boldsymbol{T}$ 的确存在,因为

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

$$f(\boldsymbol{B}\boldsymbol{y}_0, \boldsymbol{x}_0 + \boldsymbol{y}_0) = f(\boldsymbol{B}\boldsymbol{y}_0, \boldsymbol{x}_0) = f(\boldsymbol{y}_0, \boldsymbol{x}_0) \neq 0$$

于是,变换 $\boldsymbol{TB} = \boldsymbol{C}$ 把 $\boldsymbol{y}_0$ 变成 $\boldsymbol{x}_0 + \boldsymbol{y}_0$,且

$$f(\boldsymbol{y}_0, \boldsymbol{C}\boldsymbol{y}_0) = f(\boldsymbol{y}_0, \boldsymbol{x}_0 + \boldsymbol{y}_0) = f(\boldsymbol{y}_0, \boldsymbol{x}_0) \neq 0$$

而且 $\boldsymbol{C}$ 使 $\boldsymbol{x}_0$ 不变,因为 $\boldsymbol{x}_0$ 属于横截 $\boldsymbol{T}$ 的不变超平面,后者是由于

$$f(\boldsymbol{x}_0, \boldsymbol{B}\boldsymbol{y}_0 - \boldsymbol{x}_0 - \boldsymbol{y}_0) = f(\boldsymbol{x}_0, \boldsymbol{B}\boldsymbol{y}_0) - f(\boldsymbol{x}_0, \boldsymbol{y}_0) = 0$$

所以两个向量 $\boldsymbol{x}_0$ 与 $\boldsymbol{y}_0$ 使得

$$f(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0, f(\boldsymbol{y}_0, \boldsymbol{C}\boldsymbol{y}_0) \neq 0$$

这是应用引理 3 的条件,由此有结论:

引理　设 $\boldsymbol{A}$ 是任一辛变换,向量 $\boldsymbol{x}_0$ 和 $\boldsymbol{y}_0$ 满足 $f(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0$,则存在辛横截的积变换,再乘以 $\boldsymbol{A}$ 就使 $\boldsymbol{x}_0$ 和 $\boldsymbol{y}_0$ 都保持不变.

从这个结论,推证定理 4 如下.设 P 是由 $\boldsymbol{x}_0$ 与 $\boldsymbol{y}_0$ 确定的平面,$\boldsymbol{x}_0$ 与 $\boldsymbol{y}_0$ 对 $\boldsymbol{C}$ 不变且 $f(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0$.此平面 P 不是迷向子空间,它的共轭子空间 P^* 是它的互补子空间,但 P^* 的维数是 $2m-2$.变换 $\boldsymbol{C}$ 使 P^* 不变,它在 P^* 上的限制是我们前面讨论过的辛变换.如果这限制是辛横截的积,那么因为子空间的每个辛横截是 E 的辛横截的限制,故定理得证.但通过关于 m 的归纳法,可归结为 2 维空间,这时引理成立.定理证毕.

作为定理 4 的推论,每个辛变换的行列式都为 1.因此,辛群 $SP(n, K)$ 是特殊线性群 $SL(n, K)$ 的子群.

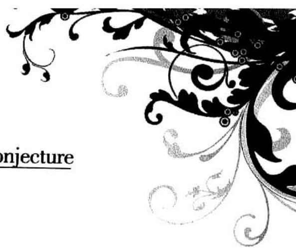
10　对称群

n 个元素的有限集合的全体置换所成的群称为对称群 S_n.这是具有 $n!$ 个元素的有限群.它可用 K^n 的线性变换实现;事实上,例如,考虑三个元素的集合 (x_1, x_2, x_3),置换 $\{x_1, x_2, x_3\} \rightarrow \{x_3, x_2, x_1\}$ 可表示成

$$\begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

所以,其矩阵表示的每行、每列都只有一个非零元素,即是 1.这对任意 n 也成立.由此,群 S_n 可视为正交群的子群;这个群的旋转有 $\frac{n!}{2}$ 个.相当于偶置换,它们构成 S_n 的子群:交代群 A_n.





第七讲　射影空间①

“数学专修班”的优秀生都能正确使用实射影空间,甚至复射影空间.但是,他第一次跟这些概念打交道,几乎总是很苦恼;这种苦恼劲,如果拿我的个人经验来说,可能持续很久.本文的目的与其说是要长篇大论射影空间的技巧,不如说是想把射影空间及其来源空间的关系搞清楚,从而弄清苦恼的根源,有助于消除它.

无穷远元素的引进遭遇到一些困难,这是从自然数直到复数相继扩充数的概念时从未出现过的.第一个困难是:这些扩充的每一个都可以使我们得到被创造出来的那些元素的新性质,而丝毫也不会失掉已有元素的性质(从实数过渡到复数失去有序体的结构是个例外);反之,引进一个(或两个)无穷远元素就几乎整个地毁灭了实数直线的代数结构.对于这一点,在初等课程里表达得过于简单了,宣称“无穷大不是一个数!”第二个困难涉及扩充的唯一性问题,对于扩充数的概念,这一点是有保证的(确切地说,例如有下述事实:包含自然数的极小阿贝尔群(就加法而言)同构于整数群),而对无穷远元素而言,唯一性问题并无保证.缺乏严格的基础使学生在含糊的问题面前无言以对,例如下面的问题:复平面和解析几何的平面在有限距离内是重合的,那怎么会一个有无穷远点,而另一个却有无穷远直线呢?

1　射影空间的定义

a) 利用把一张平面映成另一张平面的透视映射定义射影空间.设 P 与 P' 是两张不平行的平面,O 是一透视中心.有可能出现一些讨厌的例外:P 上有一条直线,它的点 M 没有象点;P' 上有一条直线,它的点 M' 不是 P' 的点的象.然而,就其余的点而言,在 P 的点 M,P' 的点 M' 以及直线 OMM' 这三者之间存在对应关系,只要知道其中任何一项,另外两项便唯一确定.但是在上述两个例外情形中,点 M 或 M' 中有一个是存在的,而从点 O 引出的直线也总是确定的.因此,如果不再着眼于 P 或 P' 的点,而着眼于 $\mathbf{R}^3$ 中从 O 引出的直线,就可以排除上述例外情形.为此,我们给出下述定义:二维射影空间 P_2 是三维欧氏空间 $\mathbf{R}^3$

① 原作者 A. Revuz(普瓦蒂埃理学院教授).

中从一点 O 引的所有直线的集合. 射影空间 P_2 中的直线就是 $\mathbf{R}^3$ 中过点 O 的平面.

一般定义:假设给了任何一个体 K(可换或不可换) 以及左向量空间 K^{n+1},它是 $n+1$ 个 K 的积空间,与 K 上的任何 $n+1$ 维左①向量空间同构. 所谓 K 上的 n 维左射影空间,记为 $P_n(K)$,是指 K^{n+1} 中所有齐次直线的集合,即 K^{n+1} 中所有一维子空间的集合.

b) 利用齐次坐标定义射影空间. 现在让我们主要从解析的观点来讨论. 设 K_{n+1}^* 是 K^{n+1} 挖掉原点后的空间(指标的位置在下表示 K_{n+1}^* 不是维数较低的同类空间的积). 在 K_{n+1}^* 中引入下面的等价关系 Δ_n. K_{n+1}^* 中的(非零) 向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 等价,如果存在非零的 $t \in K$,使得 $\boldsymbol{y} = t\boldsymbol{x}$.

相应的等价类就是 K^{n+1} 中除去原点后的齐次直线;因此,$P_n(K)$ 与商空间 K_{n+1}^*/Δ_n 之间存在一一对应,所以可以把两者看成一样.

K_{n+1}^* 的元素由其坐标 $x_0, x_1, \cdots, x_n$ 给定,所以 $P_n(K)$ 的元素由 $n+1$ 个元素 $x_0, x_1, \cdots, x_n$ 给出,这是射影空间中点的一组"齐次坐标,除了一个因子不计外是完全确定的."

2 射影空间的结构

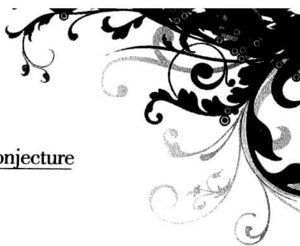
原来的向量空间结构经过这样处理以后还剩下什么呢?如果我们从一个群出发,考虑关于这个群的合成法则是正规的等价关系,那么所得到商结构仍然是一个群:商群. 但是在射影空间的情形,我们的出发点是一个向量空间,去掉了群运算的中性元素,而且,等价关系关于群的合成法则也不是正规的($\boldsymbol{x}'$,$\boldsymbol{y}$ 等价于 $\boldsymbol{y}'$ 并不蕴含 $\boldsymbol{x}+\boldsymbol{y}$ 等价于 $\boldsymbol{x}'+\boldsymbol{y}'$). 至于纯量乘法,由于它使等价类保持不变,所以并不产生射影空间中相应的运算.

那么还剩下什么呢?让我们想想讨论的出发点,想想把平面映成平面的透视映射吧. 剩下的东西不过就是共线的点罢了. 在更一般的空间中,我们指出,若 K_{n+1}^* 的 p 个元素线性无关,即 $\lambda_1\boldsymbol{x}_1 + \cdots + \lambda_p\boldsymbol{x}_p = \mathbf{0}$ 蕴含 $\lambda_1 = \lambda_2 = \cdots = \lambda_p = 0$,则与 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p$ 分别等价的 p 个元素 $\boldsymbol{x}'_1, \cdots, \boldsymbol{x}'_p$ 分别等价的 p 个元素 $\boldsymbol{x}'_1, \cdots, \boldsymbol{x}'_p$ 显然也线性无关. 相应的诸等价类叫做 P_n 的射影无关点.

射影线性簇 设 φ 是把 K_{n+1}^* 映成 P_n 的标准映射:对于 K_{n+1}^* 的每一元素 $\boldsymbol{x}$,相应的 $\varphi(\boldsymbol{x})$ 是它的模 Δ_n 类. 我们总是把 K_{n+1}^* 中的"线性对象"关于 φ 的象定义为"射影对象",例如,p 维射影线性簇就是 K_{n+1}^* 的 p 维子空间 V_p 的标准象 $\varphi(V_p)$. 由此推出,若 P_n 的 p 个点射影无关,则含有它的最小射影线性簇是 $p-$

① 或右向量空间.





1 维的.

射影空间的公理化　射影空间中射影无关的概念恰好是可以从它出发构造射影空间理论的唯一概念. 这一点 Lesieur 给出了一个极好的表述,此外,他还用序结构的语言定义了无关性①.

上面的说明基于下述事实:$P_3(R)$ 的所有射影线性簇,例如空集与全空间 $P_3(R)$ 都算在内,按包含关系确定的序关系,构成一个格,这就是说:给了两个簇 A 与 B,则所有含于 A 与 B 的簇中存在一个最大的簇,记为 $\inf(A,B)$,或 $A\cap B$,称为 A 与 B 的交②;所以包含 A 与 B 的簇中存在一个最小的簇,记为 $\sup(A,B)$ 或 $A\cup B$,称为 A 与 B 的和. 格有最小元素,即是空集,也有最大元素,即是全空间.

于是,Lesieur 提出如下公理:

Ⅰ 射影空间的线性簇构成一个格,这个格有最小元素 0 与最大元素 1. 序关系记为 $A\leqslant B$,可以读作:A 囿于 B,或 B 超过 A. 盖住 O 的元素称为点③.

Ⅱ 若点 P 不囿于簇 B,则 $B\cup P$ 盖住 B.

Ⅲ 任何一个簇 $A\neq 0$ 所超过的点的集合非空,并且 A 是其中有限多个点的和.

于是,无关性的概念表述成:$n+1$ 个点是相关的,如果其中至少有一个点囿于其余 n 个点的和中. 相反的情形则称为无关.

我们指出,如果无关点的和是一个簇,则点的个数只依赖于这个簇,而不依赖于所选的无关点,即使有很多可能的选择. 这个簇的维数就是无关点的个数减一.

若 n 是全空间(格的最大元素)的维数,则 $n-1$ 维簇称为超平面.

于是有下面第四个公理:

Ⅳ 若 A 不囿于超平面 H,则 A 盖住 $A\cap H$. 这是 Ⅱ 的对偶公理(若颠倒格的序关系,就是 Ⅱ).

这四条公理有一个重要的推论,就是簇的维数满足下述恒等式

$$d(A\cup B)+d(A\cap B)=d(A)+d(B)$$

正是从这些最简单的概念出发,使理论逐步充实以新的公理,就能重新建立起古典的射影几何,甚至仿射几何.

现在我们要抛开公理的观点. 提出这种观点只是为了强调射影空间理论的实质,以及这种观点产生的理论的优美.

---

① L. Dubreil Jacotin, L. Lesieur, R. Croisot,《Théorie des Treillis》, Paris, 1953, pp249 ~ 371.

② "交"原则上是指格论中的交,不过现在和集论中的交一致,只要把簇视为它所包含的点的集合即可. 但"和"则是格论中的运算,在这里是与集论中的并集不同的.

③ 元素 a 盖住另一个元素 b,如果 a 严格大于 b,并且 a 与 b 之间不存在任何元素.

3 射影映射①

设有同一个体 K 上的两个映影空间 P_m 与 P_n，维数分别为 m 与 n，它是从除去了一点的向量空间 K_{m+1}^* 与 K_{n+1}^* 出发定义的. 又设 φ 与 ψ 是标准映射，分别把 K_{m+1}^* 映成 P_m，把 K_{n+1}^* 映成 P_n. 对于把 K^{m+1} 映入 K^{n+1}($m \leqslant n$) 的任何单值线性映射 f，我们可以"求商"相应地得到一个射影映射 g，把 P_m 映入 P_n，g 由下列等式定义

$$g \circ \varphi = \psi \circ f$$

这些映射适合

$$\begin{array}{ccc} K_{m+1}^* & \xrightarrow{f} & K_{n+1}^* \\ \varphi\downarrow & & \downarrow\Psi \\ P_m & \xrightarrow{g} & P_n \end{array}$$

P_m 的一点 $\boldsymbol{x}$ 是 K_{m+1}^* 的一条"直线"对于 φ 的象，该直线对于 f 的象是 K_{n+1}^* 的直线，它对于 ψ 的象是 P_n 的点，按定义此点就是 $\boldsymbol{x}$ 对于 g 的象. P_m 的 p 维射影簇是 K_{m+1}^* 的 p 维(挖去 O 的)向量子空间对于 φ 的象，由此推出：射影簇对于射影映射的象是同样维数的射影簇.

推广：半线性映射. 可以将上述理论推广到一般情形，其中射影空间是定义在两个不同但却同构的体 K 与 K' 上. 设 σ 是 K 与 K' 之间的同构映射，λ^σ 是 $\lambda \in K$ 在 K' 中的象. 于是，所考虑的 f 就是与同构映射 σ 相应的半线性映射，它把 K 上的右向量空间 E 映入 K' 上的右向量空间 E'，即对任何 $\boldsymbol{x}, \boldsymbol{y} \in E$ 和 $\lambda \in K$ 有

$$f(\boldsymbol{x} + \boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}), f(\boldsymbol{x}\lambda) = f(\boldsymbol{x})\lambda^\sigma$$

如果 f 是一一的，则如上求商由 f 导出的映射称为相应的射影空间之间的射影映射.

现在我们可以在最一般的形式下陈述并证明射影几何的基本定理.

定理 设 E 与 E' 分别是体 K 与 K' 上的两个右向量空间，具有相同的维数 n；P 与 P' 是相应的射影空间. 如果存在把 P 映成 P' 的一一映射 g，使得 P 的任意共线三点的象是 P' 的共线三点，又若 $n \geqslant 3$，则有：

1) K 与 K' 同构(用 σ 表示同构映射)；

2) g 是射影映射，由一个与 σ 相应的把 E 映成 E' 的一一半线性映射 f 求商而得.

注意，定理的结论有力地确认了前面提出的观点：射影空间理论的实质包

---

① 这里简称为射影映射的概念，就是通常所说的射影线性映射. 因为我们只考虑这种映射，故不致混淆.





含在无关点的概念中.

证明① 我们用 xK 表示 P 的点,它是 E 的点 $\boldsymbol{x}$ 的标准象(对于 E',P',K' 有类似的记号).

1° P 维簇 V 的象 $g(V)$ 是 p 维簇.设 $\boldsymbol{a}_iK(1\leqslant i\leqslant p+1)$ 是生成 V 的射影无关点组;由于共线点的象仍共线,所以 $g(V)$ 含于 $g(\boldsymbol{a}_iK)$ 生成的簇 V' 内.但是,另一方面,由 $g(P)=P'$ 推出,无关点组的象仍是无关点组:事实上,假设诸 $g(\boldsymbol{a}_iK)$ 不是无关的,则将诸 $\boldsymbol{a}_i(1\leqslant i\leqslant p+1)$ 扩充成 E 的一个基底 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_{p+1},\boldsymbol{a}_{p+2},\cdots,\boldsymbol{a}_n$,象 $g(\boldsymbol{a}_iK)(1\leqslant i\leqslant n)$ 就张成一个维数 $<n-1$ 的簇,这簇应该含有 P',因为 $g(P)=P'$,但这是荒谬的,因为 $d(P')=n-1$.于是 $g(V)=V'$,否则 V' 含有形如 $g(\boldsymbol{a}K)$ 的元素,这里 $\boldsymbol{a}K$ 与 $\boldsymbol{a}_iK(1\leqslant i\leqslant p+1)$ 射影无关.由 V 与 $\boldsymbol{a}K$ 生成的簇 V_1 有维数 $p+1$,于是它的象含于维数为 p 的 V' 中.

2° K 与 K' 同构.设$(\boldsymbol{e}_i)(1\leqslant i\leqslant n)$是 E 的一个基底,$\boldsymbol{e}_1K$,$\boldsymbol{e}_2K$ 与 $\boldsymbol{e}_nK$ 是 P 的点,$\boldsymbol{e}'_1K'$,$\boldsymbol{e}'_2K'$ 与 $\boldsymbol{e}'_nK'$ 是这些点关于 g 的象.由 1° 推得,$\boldsymbol{e}'_1$,$\boldsymbol{e}'_2$ 与 $\boldsymbol{e}'_n$ 在 E' 中线性无关.设 $\boldsymbol{D}$ 是 $\boldsymbol{e}_1K$ 与 $\boldsymbol{e}_2K$ 生成的射影直线,D' 是它的象;F 是 D 与 $\boldsymbol{e}_nK$ 生成的射影平面,F' 是它的象.F 中不在 D 上的每一点可唯一地写成$(\boldsymbol{e}_1\alpha_1+\boldsymbol{e}_2\alpha_2+\boldsymbol{e}_n)K$.因此,我们可以把 D 在 F 中的补子空间与 K 上由 $\boldsymbol{e}_1$,$\boldsymbol{e}_2$ 生成的右向量空间 L 等同(D 取作无穷远直线).F 中与 D 不同的每一条射影直线相当于 L 的一直线,F 的两条直线在 D 上相交,各被对方分成两条直线,其中一条由另一条平移而得(这两条直线称为平行直线,也可通过"位似"由一条得到另一条).D' 在 F' 中的补子空间与 $\boldsymbol{e}'_1$ 和 $\boldsymbol{e}'_2$ 生成的向量空间 L' 之间也可作类似的等同.于是,对于映射 g,相应地有一个把 L 映成 L' 的一一映射 u,把一条直线变成直线,把两条平行直线变成直线,并且 $u(0)=0$,$u(\boldsymbol{e}_1)=\boldsymbol{e}'_1$,$u(\boldsymbol{e}'_2)=\boldsymbol{e}'_2$.因此,通过在 L 中构造平行直线,由 K 的两个元素 α,β 出发,容易得到元素 $\alpha+\beta$ 与 $\alpha\beta$,它们是由 $\boldsymbol{e}_1$ 生成的直线上的点的横坐标(见图 1 与 2).由于 u 保持平行性不变,若令 $u(\boldsymbol{e}_1\xi)=\boldsymbol{e}'_1\xi^\sigma$,显然可见,映射 $\xi\to\xi^\sigma$ 是把 K 映成 K' 的一一映射,并保持和与积不变,这就是所求的同构映射.

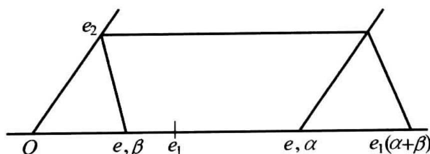


图 1

3° g 是射影映射.联结 $\boldsymbol{e}_1\xi$ 与 $\boldsymbol{e}_2\xi$ 的直线在 L 中平行于联结 $\boldsymbol{e}_1$ 与 $\boldsymbol{e}_2$ 的直线,

① 作为证明,我们引用 J. Dieudonné,《La Geometrie des Groupes Classiques》,Berlin,1955.

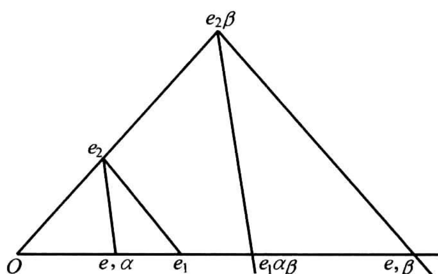


图 2

故有 $u(\boldsymbol{e}_2\xi) = \boldsymbol{e}'_2\xi^\sigma$. 另一方面,$L$ 的点 $\boldsymbol{e}_1\alpha + \boldsymbol{e}_2\beta$ 是过 $\boldsymbol{e}_1\alpha$ 与 $\boldsymbol{e}_2\beta$ 各引 $\boldsymbol{e}_2$ 与 $\boldsymbol{e}_1$ 的平行线而得的,故也有

$$u(\boldsymbol{e}_1\xi + \boldsymbol{e}_2\eta) = \boldsymbol{e}'_1\xi^\sigma + \boldsymbol{e}'_2\eta^\sigma$$

即 u 是对应于 σ 的一一半线性映射,把 L 映成 L'.

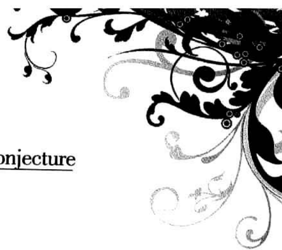
上述论证对于 P 的任何一对点 $\boldsymbol{e}_iK$, $\boldsymbol{e}_jK$ 也成立. 若用 v 表示对应于 σ 的一一半线性映射,使得 $v(\boldsymbol{e}_i) = \boldsymbol{e}'_i$, h 表示相应的把 P 映成 P' 的射影映射,则 $h^{-1}\circ g$ 是把 P 映成自身的映射,它把每一直线变成直线,并使联结两点 $\boldsymbol{e}_iK$, $\boldsymbol{e}_jK$ 的直线上的各点不变. 这是一个恒等变换,故有 $g = h$.

注 $n = 2$ 时,定理不真. 在 $K = K' = R$ 的情形下,射影变换的特征是保持两方比例不变.

4 射影空间的拓扑结构

若向量空间 E 具有拓扑结构,则求商立即得到相应射影空间的拓扑结构,实射影空间或复射影空间(即定义在实数体或复数体上的射影空间) 就是这样得到的.

作为约束,我们引述下面的重要结果: $P_n(R)$ 同胚于把对径点粘合起来的球面 S_n; $P_n(R)$ 同胚于把边界球面 S_{n-1} 的对径点粘合起来的球体 B_n; $P_n(R)$ 是紧致连通空间.



第八讲　数直线及其基本拓扑性质[1]

1 引　言

一般拓扑学是由于人们关心没有解析表达式的函数,感到有必要对连续函数给以精确的定义时应运而生的.今天,我们仍然可以说,一般拓扑学在于研究收敛与连续的概念.

实数集合 **R** 是第一个定义了这些概念的集合,不论就理论研究还是就应用而言,现在也仍然是最重要的集合.

本文要讨论的正是 **R** 的定义及其拓扑结构.

2 集合 R 是交换群,全序集,连续集

在讨论 **R** 的拓扑结构之前,我们先回顾它的基本性质;确切说,我将把 **R** 定义为具有这些基本特性的代数结构与序结构的集合.

定义 1　集合 G 是一个全序交换群,如果:

1)G 具有一个内运算,记为 $+$,使得 G 成为交换群;

2)G 具有一个全序结构,记为 $\leqslant$;

3) 上述加法与序关系是相容的,即对每个 x,$a \leqslant b$ 蕴涵 $a + x \leqslant b + x$.

显然,关系 $a \leqslant b$ 等价于 $0 \leqslant (b - a)$.

例子　全体整数的集合 **Z**;有理数集 **Q**;实系数多项式 $p(x)$ 的集合 P,其中序关系是:若 $p_2 - p_1$ 的最低次项的系数 > 0,则 $p_1 < p_2$.

定义 2　全序集 E 是连续的,如果:

1)E 的每个开区间(a,b)至少含有一点(故也含有无限多点).

2)E 的每个有上界的子界 X 都有一个上确界[2].

整数集 **Z** 满足性质 2),但不满足 1).有理数集 **Q** 正好相反.

当 **R** 已定义时,可以证明 **R** 的每个子区间是连续的.同样,令 E 是偶对

① 原作者 G. Choquet(索尔本大学教授).

② 如果对每个 $x \in X$,有 $x \leqslant a$,就说 a 是 X 的上界.如果 X 的所有上界中有一个最小的,这个最小上界称为 X 的上确界,同样定义下界与下确界的概念.

(a,b) 的集合,这里 $a\in[0,1]$,$b\in[0,1]$;定义序关系如下:

$(a,b)\leqslant(a',b')$,若 $a<a'$ 或者 $a=a'$,$b\leqslant b'$,则全序集 E 是连续的.

易知,若 E 连续,则 E 的每个有下界的子集具有一个下确界,它就是 X 的下界集合的上确界.

我们将看到,在同构的意义下,定义1与2使集合具有唯一结构.准确说,我们有:

定理1 设 G 与 G' 是两个连续的全序交换群,每一个含有的元素都不止一个.则对任意的 $a\in G,a'\in G',a>0,a'\geqslant 0$,存在唯一一个把 G 映入 G' 的增表示 f①,使得 $f(a)=a'$.当 $a'\neq 0$ 时,f 是同构;当 $a'=0$ 时,对所有 x 有 $f(x)=0$.

当 $a'\leqslant 0$ 时,把“增”字改为“降”字,可得类似结果.

这个定理的证明是模仿确立指数存在的经典初等证明.

首先证明 G(以及 G')的每个区间都具有一个中心,由此可定义 G 的元素 $\left(\frac{p}{2^r}\right)a$ 的集合 A;集合 A 在 G 上处处稠密②,f 在 A 上的定义是 $f\left(\frac{p}{2^r}a\right)=\frac{p}{2^r}f(a)$.然后将 f 扩张至整个 G.

注 设 f 与 g 是把 G 映入 G' 的两个单调表示,h 是 $h(x)=f(x)+g(x)$ 定义的映射.h 显然是一表示.另一方面,如果 $h(a)\geqslant 0(a>0)$,则对 A 的每个元素 $x\geqslant 0$ 也有 $h(x)\geqslant 0$,于是对所有的 x,$h(x)\geqslant 0$.

因此,把 G 映入 G' 两个单调表示的和仍是单调表示;准确地说,这些表示构成一个群,并与 G' 同构.

表示群的存在

定理1是基本的,因为它表明,除去同构不计外,所研究的群是唯一的.不过它没有建立群的存在性;有种种方法证明这个群的存在.

例如,G 是无限十进制展开的集合,具有通常的加法;这种方法的弱点在于:必须取消从某一位开始只含 g 的展开.对二进制展开亦然.

更古典的方法是从有理数或二进制有理数所成有序交换群 $\mathbf{Q}$ 出发,利用柯西序列方法或戴德金分割方法使 $\mathbf{Q}$ 完备化.

下面介绍一种比较方便的分割方法,排除了存在三类分割的麻烦.

序加群 $\mathbf{Q}$ 的任何非空子集 s 称为 $\mathbf{Q}$ 的起端,如果:

1) $x<y$ 和 $y\in s$ 蕴涵 $x\in s$;

2) s 不含最大元素,$s\neq\mathbf{Q}$.

例如,对于每个 $a\in\mathbf{Q}$,满足 $x<a$ 的集合 s_a 就是一个起端.

① 即 $f(x+y)=f(x)+f(y)$,$(x\leqslant y)\Rightarrow(f(x)\leqslant y)$.
② 即 G 的每个非空开区间都含有 A 的点.





设 S 是所有起端的集合;包含关系在 S 中定义一全序关系,这关系是连续的,因为任意多个起端的并集也是一个起端,否则就等于 $\mathbf{Q}$.

在 S 上定义加法如下:

$s_1 + s_2 = (x_1 + x_2)$ 的集合,这里 $x_1 \in s_1, x_2 \in s_2$.

于是 S 在上面的序关系与加法下构成一个连续的全序交换群,含有一个处处稠密的子群,这个子群同构于 $\mathbf{Q}$,同构对应是 $a \to s_a(a \in \mathbf{Q})$,一般就将这子群与 $\mathbf{Q}$ 等同.

我们就是要把集合 S 叫做数直线 $\mathbf{R}$.

3 R上的体结构

在刚才定义的数直线 $\mathbf{R}$ 上,现在还只有单独一个运算:加法;此外,还可在 $\mathbf{R}$ 内定义乘法.为此,只需把其处处稠密的子集 $\mathbf{Q}$ 上已有定义的乘法扩张至 $\mathbf{R}$ 即可.这种构造是经典的.

但是,我们要使用另一种方法,它基本定理 1,特别着眼于乘法的特性,把乘法视为一种运算在这点上和某些初等方法比较接近.

定义 3 集合 K 称为全序交换体,如果 K 具有交换体的结构,以及与 K 的代数结构相容的全序结构(用 $\leqslant$ 表示),即:

1) $a \leqslant b$ 蕴涵对每个 x 有 $a + x \leqslant b + x$;

2) $0 \leqslant a$ 与 $0 \leqslant b$ 蕴涵 $0 \leqslant ab$.

我们用 K_+ (相应地 K_+^*) 表示 K 中 $x \geqslant 0(x > 0)$ 的集合.

定理 2(存在性与唯一性) 设 G 是连续的全序交换群(用加法记号),则对 G 的每个元素 $e > 0$,G 都有唯一一个体结构,使得它的加法就是 G 的加法,它的单位元是 e,并且

(1) $0 \leqslant a, 0 \leqslant b$ 蕴涵 $0 \leqslant ab$.

这个体是交换的.

我们假设了 $e > 0$,因为在全序体中,每个平方 $x^2 \geqslant 0$;因此,若 e 是单位元,则 $e = e^2 > 0$.

唯一性 设 G 具有满足上面条件(1) 的体结构,用 xy 表示 x 与 y 的乘积.

设 $a > 0$ 是 G 的一个元素.由 $x \mapsto ax$ 定义的映射 f 是把群 G 映成自身的同构,不等式(1) 表明,f 是增表示;而 $f(e) = a$.因此,由定理 1,这个映射 f 由 a, e 以及 G 的群结构完全确定.

若 $a < 0$,有同样结构,因为在任何体中,$ax = -(-a)x$,故乘积 ax 由 a, e, G(及 x) 完全确定.

换言之,若在 G 上存在满足条件(1) 的体结构,则它完全由单位元的选取

确定.

存在性　剩下要证明:G 的确有这样的体结构,并且这个体是交换的.

上面的证明,启发我们把乘积 ax 定义成 x 对于把 e 变成 a 的增表示或降表示的象.确切地讲,就是令 $ax = f_a(x)$.

现在证明,有了如上乘法 G 就成为有序交换体:

1)$f_a \circ f_b$ 是单调表示.

因为,$f_a(f_b(e)) = f_a(b) = ab$,所以 $f_a \circ f_b = f_{ab}$.

由此推出 $f_a(f_b(c)) = f_{ab}(c)$,故 $a(bc) = (ab)c$,换句话讲,乘积是结合的.

2) 因 f_e 是恒等映射,故对于每个 b,有 $eb = b$;另一方面,$be = b$,所以 e 是单位元.

3) 对每个 $a \neq 0$,f_a 是把 G 映成 G 的同构映射,故存在元素 a',使得 $f_a(a') = e$,因此有 $aa' = e$,这蕴涵 $a'a = e$.因为,如果 $a'a = u$,则 $aa'a = au$,即 $a = au$,从而 $u = e$(因 f_a 是一一的).

换句话讲,a 具有逆元素.

4) 关系 $f_a(x + y) = f_a(x) + f_a(y)$ 成为

$$a(x + y) = ax + ay$$

今再证明

$$(a + a')x = ax + a'x$$

事实上,映射 $x \to (a + a')x$ 就是 $f_{a+a'}$,映射 $x \to ax + a'x$ 是两个单调表示 f_a 与 $f_{a'}$ 之和,根据上面的注,也是一个单调表示,用 g 表示.因为映射 $f_{a+a'}$ 与 g 在点 e 取同一值,故它们相同.

所以我们证明了乘法对加法的分配律.

上面证明的乘法性质表明,这个乘法在 G 上确定一个体结构.下面证明这个体是交换的.

5) 为此,只需证明映射 $x \to ax$ 与 $x \to xa$ 恒同.例如,设 $a > 0$.第一个映射是 f_a,设 φ_a 为第二个映射.已证 $(x + y)a = xa + ya$;另一方面,对每个 $x \geqslant 0$ 有 $xa \geqslant 0$,故 φ_a 为单调表示.因 $\varphi_a(e) = f_a(e)$,故得 $\varphi_a = f_a$.

6) 最后,注意,关系 $0 \leqslant a$ 与 $0 \leqslant x$ 蕴涵 $0 \leqslant ax$,于是定理证毕.

推论　设 K 与 K' 是两个连续的全序交换体,则存在把 K 映成 K' 的唯一的增同构映射.

这个同构映射就是把有序加法群 K 映成有序加法群 K' 的同构映射,把 K 的单位元映成 K' 的单位元.

体 R　根据上面的推论,我们可以把任何连续的全序交换体叫做实数体.不过我们还是把这个名称留给使加群 $\mathbf{R}$ 的乘法单位元为 $e = 1$ 的那个体.





指数与对数 集合 $\mathbf{R}_{+}^{*}$ 对于乘法是连续的全序群,所以根据定理1,同构于加群 $\mathbf{R}$.

把 $\mathbf{R}$ 映成 $\mathbf{R}_{+}^{*}$ 并满足 $f(1)=a$ 的单调同构映射 f,称为以 a 为底的指数.逆映射 f^{-1} 是以 a 为底的对数;这两个映射当 $a>1$ 时是增映射,$a<1$ 时是降映射.

从 $\mathbf{R}$ 与 $\mathbf{R}_{+}^{*}$ 的同构推知,$\mathbf{R}_{+}^{*}$ 的每个元都是平方,这使上述推论变得很确切.事实上,设 f 是一把 K 映成 K' 的代数表示①,对每个 $x\geqslant 0$,有 $x=y^{2}$,故 $f(x)=f(y^{2})=(f(y))^{2}\geqslant 0$,所以 f 是增表示.这是一个增同构,除非对每个 x,$f(x)=0$.

特别,对把体 $\mathbf{R}$ 映入自身的任何代数表示 f,对每个 x 有 $f(x)=x$,否则对每个 x 有 $f(x)=0$.

4　开集,闭集,点的邻域

我们刚才对 $\mathbf{R}$ 的代数结构与其序结构之间的讨论,可能使人以为,弄清这关系是定义和研究 $\mathbf{R}$ 的拓扑结构所必不可少的.我们将看到,根本不是这么回事,$\mathbf{R}$ 的拓扑结构可以只从序关系出发来定义.这就让人猜测,任何全序集的拓扑结构都可以用类似方法来定义.我们可以很容易验证情况正是如此.

$\mathbf{R}$ 中的开区间是指任何形如 $(a,b)(a<b)$,$(-\infty,a)$,$(a,+\infty)$ 的集合,以及 $\mathbf{R}$ 和 $\varnothing$.这些集合的定义是熟知的,只需用到 $\mathbf{R}$ 上存在序关系.

易见,两个开区间的交仍是开区间.

定义 4 $\mathbf{R}$ 的子集 A 称为开集,如果它是一族(有限或无限多个)开区间的并集.

也可以这样定义开集 A,如果对每个 $x\in A$,存在一个包含 x 的属于 A 的开区间.

容易推出下列性质:

O_1:任何(有限或无限多个)开集的并集仍是开集.

O_2:任意有限多个开集的交集仍是开集.

O_3:$\mathbf{R}$ 与 $\varnothing$ 是开集.

性质 O_1 与 O_3 是明显的,O_2 由如下性质推出:任意有限多个开集的交集仍是开集.

从定义还可推得:每个开区间都是开集.

注 无限多个开集的交集未必是开集:例如,开区间 $\left(-\dfrac{1}{n},\dfrac{1}{n}\right)$($n=1$,

① 即不预先假设 f 是增表示.

2,…)的交集是一点集{0},这不是开集.

定义5 当$\mathbf{R}$的子集A的补集是开集时,称A是闭集.

使用如下对偶公式

$$C(\bigcup_i A_i)=\bigcap_i(CA_i),C(\bigcap_i A_i)=\bigcup_i(CA_i)$$

从性质O_1,O_2与O_3立即推得闭集的对偶性质:

F_1:任意多个闭集的交集是闭集.

F_2:任意有限多个闭集的并集是闭集.

F_3:$\varnothing$与$\mathbf{R}$是闭集.

闭集的例子 1)每个闭区间$[a,b]$是闭集;事实上,它的补集是开区间$(-\infty,a)$与$(b,+\infty)$的并集.

2)任意有限多个闭区间的并集是闭集(因此,特别有每个有限集是闭集).

3)闭集的一个著名例子是康托的三分集.用递推定义A_n:$A_1=[0,1]$;$A_2=\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]$;$A_{n+1}$是从组成$A_n$的每个闭区间中除去中间三分之一那个开区间得到的集合.每个A_n都是闭集;要求的康托三分集A就是这些A_n的交集.

容易证明,A是$[0,1]$中按照三进制计数法其展式只含有数字0与2的那些集合.

稍后,我们要用聚点的概念对闭集的特点给出一个有趣说明.

定义6(邻域的定义) x是$\mathbf{R}$的一点,$\mathbf{R}$的子集V称为x的邻域,如果V含有一个包含x的开集.

换句话说,如果V包含一个含有x的开区间,V就是x的邻域.

例如,任何开集A都是其中每一点的邻域.反之,任何集合A,如果是其中每一点的邻域,它就是开区间的并集,所以是开集.因此开集与邻域的概念是互通的.

下面是$\mathbf{R}$中点的邻域的简单性质:

V_1:x的每一邻域含有x.

V_2:包含x的一个邻域的任一集合仍是x的一个邻域.

V_3:x的两个邻域的交集仍是x的一个邻域.

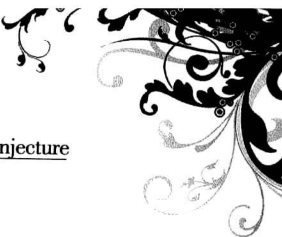
V_4:x的每个邻域V都包含x的另一邻域W,使得W的每一点以V为邻域.

V_5:$\mathbf{R}$的两个相异点a与a'有两个互不相交的邻域V与V'.

性质V_1,V_2,V_3是明显的.为证明V_4,只需取含有x并含于V的开集作为W即可.

为证明V_5,设$a<a'$,b在a与a'之间.于是只需取$V=(-\infty,b)$,与$V'=(b,+\infty)$即可.证明直线上的点列至多只能收敛于一个点,正是利用了





性质 V_5.

集合的接触点与聚点

设 A 是 $\mathbf{R}$ 的子集，$x \in \mathbf{R}$.

如果 x 的每个邻域至少含有 A 的一个点，就说 x 是 A 的接触点.

如果 x 的每个邻域至少含有 A 的一个异于 x 的点，就说 x 是 A 的聚点.

换句话讲，x 是 A 的接触点相当于 $x \in A$ 或者 x 是 A 的聚点.

容易验明，若 x 是 A 的聚点，则 x 的每个邻域含有 A 的无限多个点.

A 的一个点若非 A 的聚点，就叫做 A 的孤立点.

例子 1) 整数集合 $\mathbf{Z}$ 没有聚点.

2) 点 $\frac{1}{n}$ $(n \in \mathbf{Z}^*)$[①]所成的集合 A 有唯一的聚点 O. A 的每一点都是孤立点.

3) 有理数集没有孤立点. $\mathbf{R}$ 的每一点都是 $\mathbf{Q}$ 的聚点.

附贴包，闭包

A 的接触点的集合称为 A 的附贴包.

包含 A 的所有闭集的交集称为 A 的闭包，亦即包含 A 的最小闭集.

容易验明，这两个集合是相同的，都用 $\bar{A}$ 表示，并且相应的名称不加区别.

A 是闭集相当于 $A = \bar{A}$.

有上界闭集的上确界

从 $\mathbf{R}$ 的定义推知，$\mathbf{R}$ 的子集 A 若有上界，就有上确界 a，用 $\sup A$ 表示.

$a = \sup A$ 的特性是：$(a, +\infty)$ 不含 A 的点，并且 a 的每个邻域至少含有 A 的一个点. 特别，a 是 A 的接触点，所以要么 $a \in A$，要么 a 是 A 的聚点.

当 A 是闭集时，$A = \bar{A}$，故 $a \in A$；换句话讲，每个有上界的闭集含有它的上确界.

5 有界闭区间的紧致性

有界闭区间的一个最重要属性，就是下面的波莱尔－勒贝格定理表达的性质，它使我们可以把研究任意开覆盖的问题化为研究它的有限子开覆盖的问题，从而由局部性质得出整体性质.

定理 3(波莱尔－勒贝格定理) 设 $(w_i)_{i \in I}$ 是有界闭区间 $[a, b]$ 的一个开覆盖，即是 $\mathbf{R}$ 的一族开集，它们的并集包含 $[a, b]$，则存在一有限子覆盖 $(w_i)_{i \in J}$，它们的并集同样包含 $[a, b]$.

为证明起见，我们用 A 表示 $[a, b]$ 中使得 $[a, x]$ 能被有限多个 w_i 所覆盖

① $\mathbf{Z}^*$ 表示非零整数的集合.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

的那些 x 所成的集合；容易证明 A 含有它的上确界，并且这上确界只能是 b. 而关系 $b \in A$ 相当于定理的结论.

定理 4(波尔查诺 - 魏尔斯特拉斯定理)　有界闭区间 $[a,b]$ 的任何无限子集 X 在 $[a,b]$ 上至少有一聚点.

这是前面定理的直接推论：

若 $[a,b]$ 的任何一点 x 都不是 X 的聚点，则每个 x 都具有一个开邻域 $V(x)$，至多含有 X 的一个点，即 x 自身. 这些 $V(x)$ 构成 $[a,b]$ 的开覆盖；根据定理 3，存在有限多个这样的 $V(x)$，比如 $V(x_p)(p=1,2,\cdots,n)$ 组成 $[a,b]$ 的一个覆盖，所以 X 至多含有 n 个点 x_p.

从上面两个定理出发，容易证明把 $[a,b]$ 换成它的任意闭子集而得的相应定理.

6　连续函数，序列的极限

含糊地讲，函数 f 在一点 x_0 处连续，是指 x 与 x_0 充分接近时，$f(x)$ 与 $f(x_0)$ 就接近. 因此，可以期望，邻域的概念将对连续概念给出一个方便的定义.

下面我们只考虑定义在 $\mathbf{R}$ 的一子集 X 上并在 $\mathbf{R}$ 中取值的函数，亦即 X 上的数值函数.

定义 7　设 f 是把 $\mathbf{R}$ 的子集 X 映入 $\mathbf{R}$ 的映射，$x_0 \in X$，如果对于 $f(x_0)$ 的每个邻域 W，存在 x_0 的一个邻域 V，使得

$$(x \in V \cap X) \Rightarrow (f(x) \in W)$$

则称 f 在 x_0 连续.

如果我们规定：x_0 关于 X 的邻域是指 X 中任何形如 $(V \cap X)$ 的子集，V 是 x_0 在 $\mathbf{R}$ 中的邻域，那么，上述连续性定义中的条件相当于：对每个 W，集合 $f^{-1}(W)$ 是 x_0 关于 X 的邻域.

为简单计，现在假设 $X=\mathbf{R}$，即设 f 是把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的映射.

若 f 在每一点连续，则对 $\mathbf{R}$ 的每个开集 w，根据刚才的解释，集合 $f^{-1}(w)$ 是其中每一点的邻域，所以是开集. 反之，若对任意的开集 w，集合 $f^{-1}(w)$ 都是开集，则 f 在每一点连续. 这是连续函数的一个简要特性.

例如用上述特性，可立即证明：若 f 与 g 是把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的两个连续映射，则它们的复合 $f \circ g = f(g)$ 也连续(连续的传递性).

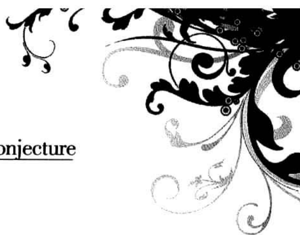
下面是一个等价特性：

把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的映射 f 是连续的，如果对 $\mathbf{R}$ 的每个闭集 F，$f^{-1}(F)$ 是闭集.

利用下列公式立即得到证明

$$f^{-1}(CA) = C(f^{-1}(A))$$





序列的极限

定义 8 设有 **R** 的点序列(x_n) 和 **R** 的一点 l. 如果对 l 的每个邻域 V,除了有限个 n 的值外,均有 $x_n \in V$,就说序列(x_n) 收敛于 l,或者说以 l 为极限.

若令 $X_n = \{x_n, x_{n+1}, \cdots\}$,则上述条件化为:至少有一个 n 使得 $X_n \subset V$.

由公理 V_5 容易证明:一个序列至多只能有一个极限.

容易证明下列结构:

1) 任何有上界的增序列(以及有下界的降序列) 都有极限.

2) 设 f 是把 X 映入 **R** 的映射,在 X 的一点 a 处连续,则对 X 中任何收敛于 a 的序列(x_n),序列($f(x_n)$) 收敛于 $f(a)$.

序列的接触点

实数序列 $a_n = \left(\frac{1}{n} - 1\right)^n$ 不收敛,但 ±1 往往可视为广义的极限. 对这个概念的分析就产生下列定义:

定义 9 设(a_n)($n = 1,2,\cdots$) 是 **R** 的点列,a 是 **R** 的一点. 如果对 a 的每个邻域 V,都存在任意大的足标 n,使得 $a_n \in V$,就说 a 是序列(a_n) 的接触点.

若用 A_n 表示集合$\{a_n, a_{n+1}, \cdots\}$,则当 a 属于每个 A_n 的附贴包时,a 就是序列(a_n) 的接触点.

因此,序列的接触点的集合是 $A = \bigcap_{n=1}^{\infty} \bar{A}_n$.

这个集合是闭的;它可能是空的,例如 $a_n = n$. 然而,使用有界闭区间的紧致性,容易证明,每一有界序列至少有一接触点.

当序列(a_n) 收敛于 l 时,则它有唯一的接触点 l. 但反之不真;例如,序列

$$\left(\frac{1}{2}, 2, \frac{1}{3}, 3, \cdots, \frac{1}{n}, n, \cdots\right)$$

有唯一的接触点 O,可是这序列并不收敛于 O. 不过,容易证明,若有界序列有唯一接触点 a,则此序列收敛于 a;这本质上是由于有界闭区间的紧致性.

R 的可数基及其应用

迄今所论拓扑性质只用到 **R** 的序结构,这些性质对于任何全序集也成立,只要其中任何有上界的子集都有上确界.

R 含有处处稠密的可数子集(即 **Q**),利用这一事实,我们可以得到与可数性有关的一些新性质. 下面举出几个. 它们的证明都是容易的:

1)**R** 中有可数多个开集(W_n),使得 **R** 的每个开集都是其中某些开集的并集(例如,可以取区间(a, b) 作为开集 W_n, $a, b \in$ **Q**). 这样的一族开集(W_n) 称为 **R** 的可数基.

2)**R** 的子集 X 的任何开覆盖都有一个有限或可数的子覆盖.

特别，两两互不相交的一族开集是有限的或可数的；由此可知，**R**的每个开集 w 都是至多可数个两两互不相交的开区间的并集．这些开区间是 w 上的下述等价关系的等价类：$x \sim y \Leftrightarrow [x, y] \subset w$.

3）若 a 是 X 的一个聚点，则 a 是 X 中由相异点组成的一个序列的极限．

4）若 a 是序列 (x_n) 的接触点，则存在 (x_n) 的一子序列收敛于 a.

5）设 f 是把 **R** 的子集 X 映入 **R** 的映射，$a \in X$．要使 f 在点 a 连续，充要条件是：对于 X 中收敛于 a 的任何序列 (x_n)，序列 $f(x_n)$ 收敛于 $f(a)$.

上述命题补足了研究序列极限时提出的一个命题．

从上述第一个性质容易推出其余性质．

7 R的一致结构

迄今我们还不曾用到 **R** 的代数结构．我们将看到，**R** 的加群结构可以确立一些新的性质；事实上，**R** 上存在平移使我们可以谈集合的微小级．

对于每个 $X \subset \mathbf{R}$, $a \in \mathbf{R}$，用 $X + a$ 表示元素 $(x + a)$ 的集合，这时 $x \in X$，并称 $X + a$ 是 X 经过平移 a 后的集合．

于是可以给出下面定义：设有 O 的邻域 V 与 **R** 的子集 X，如果可以把 X 平移到 V 中，就说 X 的微小级是 V.

这个定义容易推广至具有适当拓扑结构的群．不过，利用 **R** 是全序群这一事实，还可将上述定义精确陈述如下：

对每个 $x \in \mathbf{R}$，令 $|x| = \sup\{x, -x\}$，易证 $|a + b| \leqslant |a| + |b|$；再令 $d(x, y) = (x, y)$ 的距离 $= |x - y|$，易证：

1）若 $x \neq y$，则 $d(x, y) > 0$；$d(x, x) = 0$.

2）$d(x, y) = d(y, x)$.

3）$d(x, y) \leqslant d(x, z) + d(z, y)$（三角形不等式）.

在研究直线的一致结构时，这三个性质是基本的，经常要用到．

集合的直径

对于每个 $X \subset \mathbf{R}$，设 X' 是数 $|x - y|$ 的集合，这里 $x, y \in X$．若 X' 有上界，我们令 $\delta(X) = \sup X'$，否则令 $\delta(X) = +\infty$．$\delta(X)$ 称为 X 的直径．显然 $\delta(X) = \delta(\bar{X})$.

定义 10 设 $\varepsilon > 0$，如果 $\delta(X) \leqslant \varepsilon$，就说集合 X 的微小级是 ε.

若集合 X 与 Y 的微小级都是 ε，且 $X \cap Y \neq \varnothing$，则集合 $X \cup Y$ 的微小级是 2ε.

柯西序列

设 (a_n) 是收敛序列；若令 $A_n = \{a_n, a_{n+1}, \cdots\}$，则 $\delta(A_n) \to 0$.

我们说，序列 (a_n) 是柯西序列，如果 $\delta(A_n) \to 0$（就是说当 $p, q \to \infty$ 时，





$|a_p - a_q| \to 0$).

我们刚才已经指出,每一收敛序列都是柯西序列.逆命题则是直线的一个基本性质:

定理 5 **R** 的每一柯西序列都是收敛的.

因柯西序列有界,故至少有一接触点,又柯西序列至多只有一个接触点,故定理成立.

更经典的证明是归结为单调序列的情形.

一致连续

这个概念不能从纯拓扑的概念,即从开集或邻域出发予以定义.还要基于 R 的群结构或者距离的概念;我们要对一致连续的概念给出两个等价的定义来澄清这点:

定义 11 设 f 是把 **R** 的子集 X 映入 **R** 的映射,我们说,f 是一致连续的,如果对 O 的每个邻域 W,存在 O 的邻域 V,使得对于微小级为 V 的任何 $A \subset X$,象 $f(A)$ 的微小级是 W.

定义 11′ 记号同上.f 称为一致连续,如果 $\delta(A) \to 0$ 蕴涵 $\delta(f(A)) \to 0$.(换言之,对每个 $\varepsilon > 0$,存在 $\eta > 0$,使得 $(\delta(A) < \eta) \Rightarrow (\delta(f(A)) < \varepsilon)$)

显然,每个一致连续函数是连续的;反之不真.

然而,有一个基本定理说:当 X 是有界闭集时,把 X 映入 **R** 的任何连续映射都是一致连续的.这个结论容易从定理 3,定理 4 或者它们在有界闭集上的推广推得.

函数序列的一致收敛

这是有关值域空间的概念,对于定义空间的拓扑结构和一致结构未做任何假设.准确讲,我们有:

定义 12 设 E 是完全任意的集合,(f_n) 是把 E 映入 **R** 的一映射序列,f 是把 E 映入 **R** 的映射.我们说,序列 (f_n) 一致收敛于 f,如果对于每个 $\varepsilon > 0$,存在一个整数 $n(\varepsilon)$,使得对每个 $x \in E$ 与每个 $n \geqslant n(\varepsilon)$,我们有

$$|f_n(x) - f(x)| < \varepsilon$$

即

$$d[f_n(x), f(x)] < \varepsilon$$

在稍后的一讲中,还要更详细地研究这个概念.这里简单提醒一下,一致收敛的概念比逐点收敛要求更强些.例如,设 f_n 是定义在 $[0,1]$ 上的数值函数

$$f_n(x) = \begin{cases} n^2x(1 - nx), & \text{当 } x \in \left[0, \frac{1}{n}\right] \\ 0, & \text{当 } x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

这个函数序列逐点收敛于函数 $f \equiv 0$,但不一致收敛,因为

$$\max|f_n(x)-f(x)|=\frac{n}{4}$$

当 $n\to\infty$ 时根本不趋于 0.

8 R 上的拓扑体的结构

R 的加法与乘法和它的拓扑结构与一致结构是相容的. 现在来说明相容的意义.

定理 6 下列映射连续:

把 **R** 映入 **R** 的映射 $x\to ax$;把 $\mathbf{R}^2$ 映入 **R** 的映射 $(x,y)\to x+y$;

把 $\mathbf{R}^*$ 映入 **R** 的映射 $x\to\frac{1}{x}$;把 $\mathbf{R}^2$ 映入 **R** 的映射 $(x,y)\to xy$.

下列映射一致连续:

把 **R** 映入 **R** 的映射 $x\to ax$;把 $\mathbf{R}^2$ 映入 **R** 的映射 $(x,y)\to x+y$.

自然,对于把 $\mathbf{R}^2$ 映入 **R** 的映射必须给出连续性与一致连续性的定义. 例如,取 $\mathbf{R}^2$ 中两点的欧氏距离作为距离便可得到 $\mathbf{R}^2$ 中一致连续的概念.

注意,映射 $x\to\frac{1}{x}$ 与 $(x,y)\to xy$ 非一致连续.

利用连续与一致连续的传递性,从上面定理推出另一些熟知的结果.

例如,若 f 与 g 是两个连续(或一致连续) 的数值函数,则 $f+g$ 亦然.

我们可以对定理 6 补充一个与 **R** 的全序群结构有关的结果:

映射 $x\to|x|$ 是把 **R** 映入 **R** 的一致连续映射.

于是,从公式

$$\sup(x,y)=\frac{x+y}{2}+\frac{|x-y|}{2}$$

知,把 $\mathbf{R}^2$ 映入 **R** 的映射 $(x,y)\to\sup(x,y)$ 一致连续.

同样,若 f 与 g 是两个数值连续(一致连续) 函数,则 $|f|$ 与 $\sup(f,g)$ 亦然.

把群 R 映入群 R 的连续表示

设 f 是把 **R** 映入 **R** 的连续映射,满足

$$f(x+y)=f(x)+f(y)$$

由此立得:对每个分数 $\frac{p}{q}$

$$f\left(\frac{p}{q}\right)=\frac{p}{q}f(1)$$

因此,f 在 **Q** 上的限制就是 $x\to ax$.

把 **R** 映入 **R** 的这两个映射 f 与 $x\to ax$ 都连续,并且在处处稠密的集合 **Q** 上





相同,所以(利用一个容易证明的但又重要的性质)它们是恒同的.

因此,把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的每个连续表示(对于加法)是单调的并且具有形式 $x \to ax$.

设 g 是把 $\mathbf{R}_+^*$ 映入 $\mathbf{R}_+^*$ 的映射并满足 $g(xy) = g(x)g(y)$,利用 $\mathbf{R}$ 与乘法群 $\mathbf{R}_+^*$ 之间的同构映射,立即得到关于 g 的类似结果.

$\mathbf{R}$ 的子群

本文最后讲一个简单而常用的结果:

定理 7 $\mathbf{R}$ 的任何加法子群要么是处处稠密的,要么具有形式 $a\mathbf{Z}$(点 an, $n \in \mathbf{Z}$ 的集合).

由此推知,$\mathbf{R}$ 的任何闭的加法子群要么等同于 $\mathbf{R}$,要么是群 $a\mathbf{Z}$.

看一个应用:设 f 是 $\mathbf{R}$ 上的数值函数;如果对每个 x,有 $f(x + a) = f(x)$,就说 a 是 f 的一个周期.

f 的周期所成的集合 P 显然是 $\mathbf{R}$ 的加法子群.当 f 连续时,P 是闭集.

因此,要么有 $P = \mathbf{R}$,这蕴涵 $f =$ 常数,要么 P 具有形状 $a\mathbf{Z}$,恒可假设 $a \geqslant 0$,并称这个数 a 是 f 的最小周期.

第九讲　欧氏空间与尺度空间·尺度概念与拓扑概念①

1　数直线复习

在前一讲演中,G. Choquet 提到了实数直线 $\mathbf{R}$ 的代数性质,并且立足于 $\mathbf{R}$ 的序结构定义了把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的连续映射.区间(a,b)是满足 $a<x<b$ 的实数 x 的集合,把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的映射 f 在 x_0 连续是指:对于任何含有 $f(x_0)$ 的区间(α,β),存在一个含有 x_0 的区间(a,b)使得 $x\in(a,b)$ 蕴涵 $f(x)\in(\alpha,\beta)$.也可用逆映射 f^{-1} 来定义(A 是 $\mathbf{R}$ 的子集,$f^{-1}(A)=\{x\in\mathbf{R}\mid f(x)\in A\}$).如果对每个含有 $f(x_0)$ 的区间(α,β),$f^{-1}(\alpha,\beta)$包含一个含有 x_0 的区间,就说 f 在 x_0 连续.如果用开集表示任意多个开区间的并集,那么可以指出:把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的映射在每一点处连续的充要条件是:每个开集的逆象仍是开集(空集视为开集).

我们的首要目标,是要利用数直线 $\mathbf{R}$ 的另一结构来定义把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的映射 f 的连续性:用两个实数之差的绝对值表示这两个实数的距离,可以对 f 在点 x_0 处的连续性给出三个等价的定义如下:

(1) 对于每个 $\varepsilon>0$,存在 $\eta>0$,使得 $|x-x_0|<\eta$ 蕴涵 $|f(x)-f(x_0)|<\varepsilon$.

(2) 当 x 到 x_0 的距离趋于 0 时,$f(x)$ 到 $f(x_0)$ 的距离也趋于 0.

(3) 对每个以 $f(x_0)$ 为中心的区间 I,$f^{-1}(I)$ 包含一个以 x_0 为中心的区间.

由于(3) 以及含有一点的任何开区间必包含一个以此点为中心的开区间,所以上述连续性定义(3) 与基于 $\mathbf{R}$ 的序结构的定义是等价的.这里,$\mathbf{R}$ 的开集可以定义为:开集是指一个集合 O,使得 $x\in O$ 蕴涵 O 中含有一个以 x 为中心的开区间 I.

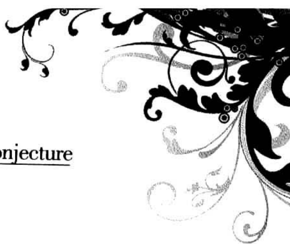
上述连续性定义(3) 的优点是适用于任何定义了距离的空间.

2　距离的定义

E 是一个集合,它的元素用 $x,y,\cdots$ 表示.E 上的距离是指把 $E\times E$ 映入 $\mathbf{R}^+$

① 原作者 A. Revuz(普瓦蒂埃理学院教授).





（非负实数集）的一个映射 d，具有下列性质：

(1) $d(x,y) = 0 \Leftrightarrow x = y$.

(2) $d(x,y) = d(y,x)$（对称）.

(3) $d(x,y) \leqslant d(x,z) + d(z,y)$（三角形不等式）.

具有这种距离的 E 叫做尺度（或距离）空间. 显然，映射 $(x,y) \to |x - y|$ 就是 $\mathbf{R}$ 上的一个距离.

设 E 与 F 是两个尺度空间，关于连续性的概念，上面的定义(2)可以原封不动地照搬. 与(1)相当的定义是：设 d 是 E 上的距离，δ 是 F 上的距离，则对每个 $\varepsilon > 0$. 存在 $\eta > 0$，使得 $d(x,x_0) < \eta \Rightarrow \delta[f(x),f(x_0)] < \varepsilon$. 与(3)相当的定义要引入球的概念：中心为 x_0，半径为 r（正数）的开球（闭球）是指满足 $d(x_0,x) < r(\leqslant r)$ 的点 x 的集合. 于是与(3)相当的定义是：中心为 $f(x_0)$ 的开球的逆象包含一个中心为 x_0 的开球.

在继续讨论从距离出发定义连续性之前，就距离本身略谈几句. 直线上，平面($\mathbf{R}^2$)中，三维空间($\mathbf{R}^3$)中那些“具体”的距离都满足距离的公理，所以就这些情形而言那几条公理似乎已经“合情合理”了. 不过考虑一下其中每一条的意义，以及弄清楚在什么情况下缺少一条就会给距离的使用造成阻碍，是不无益处的.

第一个公理表明，每个元素到它自身的距离是零，到另一个元素的距离为正. 如果我们设想一种“距离”，满足(2)与(3)，以及 $x = y \Rightarrow d(x,y) = 0$（即不要求 $d(x,y) = 0 \Rightarrow x = y$），并且的确存在若干对不同的元素，彼此距离均为零（这样的距离称为准距离）. 那么会有什么情况呢？二者必居其一：要么是因为该距离选择不当，必须考虑别的距离；要么是由于彼此距离为零的元素尽管可以说是不同的，但却不应该区别对待，因为它们在所论述的问题中起同样的作用. 在后一情形，我们注意：(1) 由 $d(x,y) = 0$ 所定义的关系 $x\rho y$ 是一等价关系；(2) 一个元素 x 到同一等价类的两个元素 y 与 z 的距离一样，这是因为 $d(z,y) = 0$ 蕴涵

$$d(x,z) \leqslant d(x,y) + d(y,z) = d(x,y)$$
$$d(x,y) \leqslant d(x,z) + d(z,y) = d(x,z)$$

所以，两个模 ρ 等价类的距离可以定义为一个等价类中任一元素到另一个等价类中任一元素的距离. 商空间 E/ρ 是一尺度空间，是把 E 中彼此距离为零的元素看成一样而得. 平面上点的横坐标差的绝对值就是上述准距离的一个简单例子：等价类是直线 $x =$ 常数，空间 E/ρ 与数直线同构（见后）.

第二个公理，即距离的对称性，尽管显得自然，但却不是随便一个具体的距离都能满足的，例如走山路按时间测行程所得的距离，上山距离就比下山距离大. 不过，碰到不对称的距离也不必大惊小怪，因为，若令

$$\delta(x,y) = \sup[d(x,y), d(y,x)]$$

那么 δ 就是满足三条公理的距离(只要 d 满足(1) 与(3)).

第三条公理的好处,直观地讲,就是保证了“邻域的邻域仍是邻域”.另一些不等式也有同样的效果,例如,φ 是正的实函数,在点(0,0) 为零且连续,则形如

$$d(x,y) \leqslant \varphi(d(x,z), d(z,y))$$

的不等式至少保证了“很接近的很接近仍很接近”.

经典形式的三角形不等式有两条优点:

(1) 简明.

(2) 适用范围很大.公立中学四年级为颁发微积分学考试合格证书而布置学习的所有空间都可以赋予尺度结构,这些结构确实有用,而不是为了考试硬塞进去的.

距离的例子①

注意,任何集合上都至少可以定义一个距离,就是令 $d(x,x) = 0$;若 $x \neq y$ 则 $d(x,y) = 1$.(这就是土地丈量员用一把可以任意伸缩的米尺可能丈量出来的距离)这显然是一个极端的例子,除了对尺度空间的某个命题提出可能的反例时能叫人一眼看穿以外,几乎不起什么作用.

在空间 $\mathbf{R}^n$ 中,对于 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$, $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)$ 我们可以定义一个距离 $d_1(\boldsymbol{x},\boldsymbol{y}) = \sqrt{\sum_i (x_i - y_i)^2}$(欧几里得距离).

公理(1) 与(2) 显然满足.要证明三角形不等式,令

$$x_i - z_i = a_i, z_i - y_i = b_i$$

则

$$x_i - y_i = a_i + b_i$$

问题就是要验明

$$\sqrt{\sum (a_i + b_i)^2} \leqslant \sqrt{\sum a_i^2} + \sqrt{\sum b_i^2}$$

此不等式等价于

$$\sum (a_i + b_i)^2 \leqslant \sum a_i^2 + \sum b_i^2 + 2\sqrt{(\sum a_i^2)(\sum b_i^2)}$$

即等价于

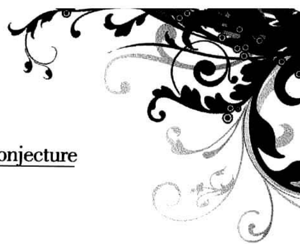
$$\sum a_i b_i \leqslant \sqrt{(\sum a_i^2)(\sum b_i^2)}$$

上面的不等式,当 $\sum a_i b_i \leqslant 0$ 时是成立的,当 $\sum a_i b_i > 0$ 时,由下面的许瓦兹不等式推出

$$(\sum a_i b_i)^2 \leqslant (\sum a_i^2)(\sum b_i^2)$$

① 除了数学上的例子以外,举一些日常生活中的距离,可能不无好处:电话谈话的单价,旅行的期限,旅行费用,……,这些例子我们虽然不习惯于称之为距离,但一般却具有所说的那些数学性质.





因三项式 $\sum(a_i + b_i\lambda)^2$ 大于或等于0,故许瓦兹不等式成立.

注意,仅当 $\frac{a_i}{b_i} = -\lambda(i = 1,2,\cdots,n)$ 时,等号才成立,即仅当

$$\frac{x_1 - z_1}{y_1 - z_1} = \cdots = \frac{x_i - z_i}{y_i - z_i} = \cdots = \frac{x_n - z_n}{y_n - z_n}$$

亦即三点 x,y,z 共线时,等号才成立.

$\mathbf{R}^n$ 中另一个方便的距离是

$$d_2(\boldsymbol{x},\boldsymbol{y}) = \sup_i |x_i - y_i|$$

还有第三个距离 $d_3(\boldsymbol{x},\boldsymbol{y}) = \sum_i |x_i - y_i|$.

若将诸坐标视为笛卡儿直角坐标,则 d_1 确定的球体的边界,在 $\mathbf{R}^2$ 中是圆周,在 $\mathbf{R}^3$ 中是球面.(如果用斜坐标就是椭圆与椭球面)在直角坐标下,d_2 确定的球体的边界,在 $\mathbf{R}^2$ 中是各边平行于坐标轴的正方形(在 $\mathbf{R}^3$ 中是立方体),而 d_3 确定的球体的边界,在 $\mathbf{R}^2$ 中是各边平行于坐标角平分线的正方形(在 $\mathbf{R}^3$ 中是正八面体).

若 d 是某个空间上的距离,f 是定义在 $\mathbf{R}^+$ 上的正值实函数,$f(0) = 0$,并且对每个 $x > 0, y > 0$ 有 $f(x + y) \leqslant f(x) + f(y)$,则 $f(d)$ 也是距离,它和 d 确定出同样一些球(一般说来,每个球对于 d 与 $f(d)$ 有不同的半径).特别,可取增凹函数作为 f.例如,若 d 是一距离,则 $\frac{d}{1 + d}$ 也是一个距离,这是一个有界距离,和原来的距离确定同一些球.

进一步,如果同一空间上定义了若干距离 $d_1, d_2, \cdots, d_p$,容易验明

$$d = \sqrt{d_1^2 + d_2^2 + \cdots + d_p^2}$$

也是一个距离(许瓦兹不等式的推论),还有 $d = d_1 + d_2 + \cdots + d_p$ 及 $d = \sup_i d_i$ 也都是距离.(在后一情形,由 d 确定的中心为 m、半径为 r 的球是各个 d_i 确定的同一中心、同一半径的 p 个球之交)

最后注意,若同一空间上定义了距离序列 (d_n),那么,由此可以派生出一个新的距离

$$d = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n}{1 + d_n}$$

测地距离

设有尺度空间 E 与 E 的子集 F,把 $E \times E$ 映入 $\mathbf{R}^+$ 的映射(距离) d 在 $F \times F$ 上的限制显然是 F 上的距离,称为由 E 的距离导出的距离.例如,若 E 是 $\mathbf{R}^3$,F 是球面 S,那么 $\mathbf{R}^3$ 的欧氏距离在 S 上导出一个距离(弦长),也称为把 S 嵌入 $\mathbf{R}^3$ 的距离.由此出发,还可定义另一个距离:测地距离.

给出 F 的两点 a 与 b，假设 F 中至少存在一条联结 a 与 b 的道路. 所谓道路是指把区间 $[0,1]$ 映入 F 的一个连续映射 f，满足 $f(0)=a, f(1)=b$. 对于任何有限序列 $\{t_i\}: t_0=0 \leqslant t_1 \leqslant \cdots \leqslant t_i \leqslant \cdots \leqslant t_n=1$，相应地得到点 $a, \cdots, m_i=f(t_i), \cdots, b$ 以及正实数 $L(\{t_i\})=d(a,m_1)+\cdots+d(m_i,m_{i+1})+\cdots+d(m_{n-1},b)$. 对于所有的有限序列 $\{t_i\}$，相应的数 L 的集合，如果有上确界，就说道路是可求长的，$\sup L(\{t_i\})=L(f)$ 就是道路 f 的长.

现在验明 $\delta(a,b)=\inf L(f)$（inf是对所有联结 a, b 的连续道路来计算的）是一个距离. 事实上：

(1) $\delta(a,b)=0 \Leftrightarrow a=b$ 是显然的. 因为若 $a=b$，则存在一条联结 a 与 b 的道路，长为零（对于每个 $t \in [0,1], f(t)=a$），若 $a \neq b$，则每条道路的长度至少等于 $d(a,b)>0$.

(2) 对称性明显.

(3) 给了 $\varepsilon>0$，则存在道路 f_{ac} 与 f_{cb}，使得

$$L(f_{ac})<\delta(a,c)+\varepsilon, L(f_{cb})<\delta(b,c)+\varepsilon$$

可用明显的方法定义两条道路的和：相应的映射 φ 定义成

$$\varphi(t)=\begin{cases} f_{ac}(2t), & \text{当 } 0 \leqslant t \leqslant \frac{1}{2} \\ f_{cb}(2t-1), & \text{当 } \frac{1}{2} \leqslant t \leqslant 1 \end{cases}$$

和的长是长的和. 因为 φ 是联结 a 与 b 的一条道路，故有

$$\delta(a,b) \leqslant L(\varphi)=L(f_{ac})+L(f_{cb})< \delta(a,c)+\delta(b,c)+2\varepsilon$$

因 ε 任意，由此推得三角形不等式. 这就是从 $\mathbf{R}^3$ 的欧氏距离出发在球面 S 上定义的测地距离. 利用上面的定义，可以初等地证明，测地距离是联结两点的两段大圆弧中最小一个的长.

3 同尺映射

把尺度空间 E（距离为 d）映成尺度空间 F（距离为 δ）的一一映射 f 叫做同尺映射，如果对于 E 的每一对元素 (a,b) 有

$$d(a,b)=\delta(f(a),f(b))$$

初等几何中所谓的位移变换和对称变换都是同尺映射，原来的名称是不恰当的（最好是叫做正同尺映射或反同尺映射）①.

① 此外，看来正是利用了距离和同尺映射这些概念，才能最容易地给出初等几何的某种简单而"具体"的公理体系，即是说这些公理表达了几何中通常的运算（测量距离、弯曲，等）.





显然,若 E' 是 E 的子空间,则 f 在 E' 上的限制是把 E' 映成 $f(E')$ 的同尺映射.但是,E 与 F 的两个子空间之间的同尺映射可以扩张成整个空间上的映射,这个断言并不总是对的,因而是某些定理要讨论的目的.

例如,若 $E = F = \mathbf{R}^2$,由不共线三点组成的两个子集之间的同尺映射可以唯一地扩张至整个空间.这是三角形三边等则全等这个命题的升华形式.若 E 与 F 是具有同样半径的两个球面,上述结论也真,这里距离是嵌入距离或相应的测地距离.就用这些距离,对于半径不同的两个球面,这个结论就不对.(这些例子表明,通常那些全等情形的证明"含有多少未曾表达出来的公理"啊)

4 连续性

现在再来讨论把尺度空间 E 映入尺度空间 F 的连续映射 f 的第三个定义:f 在一点 x_0 连续,意指每个中心为 $f(x_0)$ 的开球的逆象都包含一个中心为 x_0 的开球.这导致引入下列概念:

定义 任何集合,如果包含以 x 为中心的一个开球,就称为点 x 的邻域.

于是可以说:"f 在点 x_0 连续的充要条件是:$f(x_0)$ 的每个邻域的逆象是 x_0 的一个邻域",这一简单而普通的陈述有一个好处,就是保持连续概念的直观含义.

假设 f 在 E 的每一点连续,利用 E 和 F 的开集可以把连续的概念表达得更简单.

定义 尺度空间中的开集是任意多个开球的并集.

注意,若 $y \in B(x,r)$,则 $d(x,y) = r' < r$,所以根据三角形不等式有 $B(y,r-r') \subset B(x,r)$,由此得到开集的另一定义:开集是其中每一点的邻域.由此推出:"f 是把 E 映入 F 的连续映射(即在 E 的每一点连续)的充要条件是:F 的每个开集的逆象是 E 的开集."

用连续性的上述定义容易证明下面定理:若 E,F,G 是三个尺度空间,f 是把 E 映入 F 的连续映射,g 是把 F 映入 G 的连续映射,则把 E 映入 G 的复合映射 $g \circ f$ 仍连续.

若 $A \subset F$,CA 是 A 的补集,则

$$f^{-1}(CA) = Cf^{-1}(A)$$

因此,如果闭集是指其补集是开集的集合,那么,在上面连续性定义中,把开字换成闭字,便得到一个等价的定义.

利用集合的接触点的概念,可以得到闭集的一个内在定义.

定义 如果对于点 a 的每个邻域 V,我们有 $V \cap A \neq \varnothing$,$a$ 就是 A 的接触点.(显然只需中心为 a 的开球是这样就行了)集合 A 的附贴包 $\bar{A}$ 是 A 的所有接

触点的集合.对于任何 A,有 $A \subset \bar{A}$.

定理 一个集合是闭集的充要条件是:这个集合等于它的闭包.

事实上,A 是闭集等于说 CA 是开集.若 $a \notin A$,则存在一球 $B(a,r)$ 含于 CA,故 $B(a,r) \cap A = \varnothing$,$a \notin \bar{A}$,由此 $\bar{A} \subset A$,即 $\bar{A} = A$.反之,若 $A = \bar{A}$,则对每点 $a \notin A$,存在一个球 $B(a,r)$,使得 $B(a,r) \cap A = \varnothing$,由此 $B(a,r) \subset CA$,故 CA 是开集.

容易验明闭球是闭集(这就保证了用语的呼应).

开集有两个很重要的性质:

(1) 任意多个开集的并集是开集,这是定义的直接推论.

(2) 任意有限个开集的交集是开集.(空集算做开集)事实上,设 O_1,$O_2,\cdots,O_n$ 是 n 个开集,其交集非空,$a \in \bigcap_{i=1}^{n} O_i$.则对 $i = 1,2,\cdots,n$ 存在一个球 $B(a,r_i)$ 含于 O_i.故 $B(a,\inf_i r_i) \in \bigcap_{i=1}^{n} O_i$,明所欲证.

闭集是开集的补集,因此有对偶性质(读者也可使用闭集的定义 $A = \bar{A}$ 给以直接的证明):

(F_1) 任意多个闭集的交集是闭集.

(F_2) 任意有限个闭集的并集是闭集.

5 拓扑概念

设有两个集合 E 与 F,E 的距离 d(F 的距离 δ)使 E(F)成为尺度空间 E_d(F_δ)·$C(E_d, F_\delta)$ 表示把 E_d 映入 F_δ 的所有连续映射的集合.(这是把 E 映入 F 的所有映射所成集合的子集)这里可以提一个重要问题:d 或 δ 的改变或两者同时改变,能否不引起 $C(E_d, F_\delta)$ 的改变?下面回答一个相近的问题.

设 d_1 是 E 上的已知距离,E 上应该引进什么样的距离 d_2,才能对任何尺度空间 F_δ 有

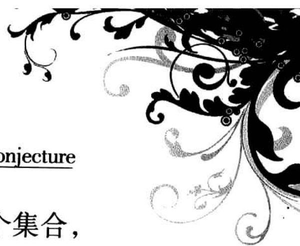
$$C(E_{d_2}, F_\delta) = C(E_{d_1}, F_\delta)$$

显然,只需与 d_1 相应的开集族 O_1 恒等于与 d_2 相应的开集族 O_2 即可.这条件还是必须的.因为如果有一集合属于 O_1 而不属于 O_2,那么只要取 $F_\delta = E_{d_1}$,就会发现把 E 映入自身的恒等映射属于 $C(E_{d_1}, E_{d_1})$ 但不属于 $C(E_{d_2}, E_{d_1})$.

d_1 与 d_2 确定同一开集族的条件与下列条件等价:对于每一 $a \in E$,关于 d_1 的每个球 $B_1(a,r)$ 包含一个关于 d_2 的球 $B_2(a,r')$,反之亦然.

据此,可以验明,前面在 $\mathbf{R}^n$ 中定义的三个距离 d_1, d_2, d_3 确定出同样的开集族,所以对于每个尺度空间 F_δ,这三个距离确定出同一集合 $C(\mathbf{R}^n, F_\delta)$.





因此,尽管给了 E 的尺度结构就能确定集合 $C(E,F)$.但是给了这个集合,甚至对任何 F 给了相应的集合,也并不能毫不含糊地确定 E 的尺度结构.

我们虽然用距离来定义了连续,但归根结底,连续性与距离本身无关,只与由距离确定的开集族有关.

因此,如果我们只关心定义在空间 E 上的映射的连续性,那么精确说明空间的尺度结构就不是很必要的了.正因为如此,我们才来定义更一般的结构:拓扑结构.

给了一集合 E,我们说,E 上定义了一个拓扑结构 $\mathfrak{T}$,如果存在 E 的一族子集 O,满足下面的两条公理:

(1)O 中任意多个集合的并集仍是 O 中的集合.

(2)O 中任意有限个集合的交集仍是 O 中的集合.

(关于 O 的空交与空并的约定①蕴含 E 及 E 的空子集属于 O)

O 中的子集合称为 E 关于拓扑结构 $\mathfrak{T}$ 的开集,E 叫做拓扑空间.

若 E 与 F 是两个拓扑空间,把 E 映入 F 的映射 f 连续,如果 F 的每个开集的逆象是 E 的开集.

前面给出的所有定义立即平移过来:开集的补集称为闭集,所以闭集族 $\mathfrak{F}$ 满足公理 F_1 与 F_2;拓扑结构 $\mathfrak{T}$ 显然可以用 E 中满足这两条公理的子集族 $\mathfrak{F}$ 来定义.

同样,一个点的邻域是一个集合,它包含一个含有这点的开集;集 A 的接触点是指它的任何邻域 V 都满足 $V \cap A \neq \varnothing$;集合 A 的闭包 $\bar{A}$ 是 A 的所有接触点的集合.可证闭集是满足 $A = \bar{A}$ 的集合.

两个拓扑空间 E 与 F 之间的一一且双方连续的映射称为同胚.

这些定义一旦确认,就可以提出许多问题,例如:

集合 E 上给出的拓扑结构能否由 E 上的一个尺度结构产生?

上面定义的那些拓扑结构的一般性质应该加以什么样的限制,才能使之具有某种特殊的性质,例如 $\mathbf{R}^n$ 的性质?

拓扑学作为一门理论,其任务就是回答这类问题.

① 约定:空交为全空间 E,空并为 E 的空子集.

第十讲　与尺度空间结构有关的概念[①]

M. Revuz 在其尺度空间的讲演中，定义了尺度空间的拓扑结构，以及把一个尺度空间映入另一个尺度空间的映射的连续性.

我在前面讲数直线的时候指出，除连续性外，还有一致连续. 我们将看到，在任意的尺度空间中，这个概念仍有意义. 更一般，我们将研究尺度空间中集合的微小级及其若干推论.

尺度空间是具有距离 $d(x,y)$ 的集合 E，这距离是把 E^2 映入 $\mathbf{R}_+$ 的映射，满足：

1) 若 $x \neq y$，则 $d(x,y) > 0$；若 $x = y$，则 $d(x,y) = 0$.

2) $d(x,y) = d(y,x)$.

3) $d(x,y) \leqslant d(x,z) + d(z,y)$.

定义 1　A 是尺度空间的子集，当 $x,y \in A$ 时，距离 $d(x,y)$ 所成集合的上确界 $\delta(A)$ 称为 A 的直径.

这个数 $\geqslant 0$(有限或无限).

例如，平面圆盘的直径就是普通意义下直径的长；等边三角形的直径是其边长；直线上区间 (a,b) 的直径是 $(b-a)$.

我们恒有 $\delta(A) = \delta(\bar{A})$[②]；仅当 A 是空集或一点集时，才有 $\delta(A) = 0$.

当 A 的直径有限时，就说 A 有界.

1　一致连续

设 f 是把尺度空间 E 映入尺度空间 F 的映射. 利用距离，f 在 E 中的连续性可表达如下

$$\forall \varepsilon > 0, \forall x \in E, \exists \eta > 0$$

使得
$$(d(x,y) < \eta) \text{ 蕴涵 } (d(f(x),f(y)) < \varepsilon)$$

① 原作者 G. Choquet (索尔本大学教授).

② $\bar{A}$ 表示 A 的闭包.





数 η 与 ε 及 x 有关;如果存在这样的 η,它与 x 无关,就说 f 一致连续,或者更简洁地说:

定义 2 如果对于每个 $\varepsilon > 0$,存在 $\eta > 0$,使得对每个 $A \subset E$,条件 $\delta(A) < \eta$ 蕴涵 $\delta(f(A)) < \varepsilon$,就说 f 一致连续.

如果把直径 $\delta(f(A))$ 称为 f 在 A 上的振幅,那么,f 的一致连续还可说成:当 A 的直径小时,f 在 A 上的振幅也小.利用连续模的概念,可以把这一说法表达精确:设 φ 是定义在$[0, +\infty)$ 的数值增函数;如果对所有的 $x, y \in E$,有

$$d(f(x), f(y)) \leqslant \varphi(d(x, y))$$

就说 f 以 φ 作为连续模.

可以验证:f 的一致连续性等价于存在连续模 φ,使得当 $u \to 0$ 时,$\varphi(u) \to 0$.

最常用的连续模是线性函数 $\varphi(u) = ku$;当 f 有这样的连续模时,就说 f 是比值为 k 的李普希兹映射.于是上面的条件成为

$$d(f(x), f(y)) \leqslant kd(x, y)$$

例子 1) 若 f 是 $\mathbf{R}$ 上的可微函数,导数 f' 满足 $|f'| \leqslant k$,则 f 是比值为 k 的李普希兹函数(用有限增量定理证明).

2) A 是 E 的子集,$x \in E$;数 $d(x, A) = \inf\limits_{y \in A} d(x, y)$ 称为 x 到 A 的距离.

把 E 映入 $\mathbf{R}$ 的映射 $x \to d(x, A)$ 是比值为 1 的李普希兹映射.

把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的映射 $x \to |x|$ 是一个特例.

把 $\mathbf{R}^2$ 映入 $\mathbf{R}$ 的映射 $(x, y) \to x + y$ 是李普希兹映射.而映射 $(x, y) \to xy$ 则不是;相反,如果规定 x 或 y 属于 $\mathbf{R}$ 中一个有界子集,这个映射则是李普希兹映射.

易见,函数 f 的一致连续蕴涵 f 在每一点连续,其逆不真;例如,函数 $x \to x^2$ 连续,但是关系

$$x^2 - y^2 = (x - y)(x + y)$$

表明,条件 $|x - y| < \eta$ 并不蕴涵 $|x^2 - y^2|$ 有上界.

下面的性质留作练习,根据这性质可以作出若干类似例子:

设 f 是一致连续的数值函数,定义在 $E = \mathbf{R}^n$ 上,或者更一般,定义在 $\mathbf{R}^n$ 的一个凸集 E 上.则存在两个常数 a 与 $b \geqslant 0$,使得对每个 $x \in E$ 有 $|f(x)| \leqslant a\|x\| + b$,这里 $\|x\|$ 表示 0 到 x 的距离.

一致连续函数的复合

1) 两个一致连续函数的复合 $g \circ f$ 仍一致连续.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

2) 若 f 与 g 分别是比值为 k_1 和 k_2 的李普希兹函数,则 $g\circ f$ 是比值为 k_1k_2 的李普希兹函数.

应用:设 f 是定义在 E 中的数值函数;若 f 一致连续(或者说比值为 k 的李普希兹函数),则 $|f|$ 亦然.据此,容易证明,若 f 与 g 是一致连续(或李普希兹函数),是 $(f+g)$, $\sup(f,g)$, $\inf(f,g)$ 亦然;只要 f 或 g 有一个还在 E 中有界, fg 亦然.

两个尺度空间的同构

我们知道,空间的拓扑结构完全由其开集确定,从开集概念出发可以得到两个空间 E 与 F 的一种同构概念.把 E 映成 F 的一一映射 f 称为拓扑同构(更经常称为同胚),如果 f 使 E 与 F 的开集互相映成,或者说如果 f 与 f^{-1} 连续.

同样,从尺度空间上一致连续的概念出发也可以得到一种同构的概念:

定义 3 把尺度空间 E 映成另一尺度空间 F 的一一映射 f 称为一致同构,如果 f 与 f^{-1} 一致连续.

拓扑同构不必是一致同构;例如,把 $\mathbf{R}$ 映成 $\mathbf{R}$ 的映射 $x\to x^3$ 是同胚,但不是一致同构.

因此,在同一个集合 E 上的两个距离 d_1 与 d_2 可以确定同样的拓扑结构,但不必确定同样的"一致结构".要使 d_1 与 d_2 在 E 上定义同样的一致结构,充要条件是:当 $d_1(x,y)$ 很小时, $d_2(x,y)$ 也很小; $d_2(x,y)$ 很小时, $d_1(x,y)$ 也很小.例如,若 d 是 E 上的距离,对于一致结构,这个距离等价于距离 $\frac{d}{1+d}$ 与 $d^{\alpha}(0<\alpha\leqslant 1)$;在 $\mathbf{R}^2$ 内,常常使用下列范数决定的距离

$$\|x\| = (x_1^2+x_2^2)^{\frac{1}{2}} \quad 或 \quad |x_1|+|x_2| \quad 或 \quad \sup(|x_1|,|x_2|)$$

(这里 x_1 与 x_2 是 x 的坐标)这些不同的距离彼此等价.

正如把 E 映入 F 的映射 f 的连续性只依赖于 E, F 的拓扑结构一样, f 的一致连续性也只依赖于 E, F 的一致结构.

定义在紧致空间上的连续映射

我们知道,有些连续函数并不一致连续,例如,把 $\mathbf{R}$ 映入 $\mathbf{R}$ 的映射 $x\to x^2$,这个例中,空间 E 与 F 不是有界的;但是 E 与 F 有界时,也存在类似的例子,例如,把 $(0,1)$ 映入 $[-1,+1]$ 的映射 $x\to\sin\frac{1}{x}$ 连续但不一致连续.

反之,当 E 紧致时,就不存在类似的反例.这就是下面的基本定理:





定理 1 把紧致尺度空间 E 映入任一尺度空间的连续映射都是一致连续的.

证明见关于紧致空间的讲演.

2 柯西序列,完备空间和逐次逼近法

尺度空间中点列收敛的概念涉及点列的极限点.然而,要知道一个序列是否收敛却不一定要求出它的极限,这一点往往是很有用的.正是柯西首先对于实数序列得出一个不依赖于极限值的收敛准则.

定义 4 设 E 是尺度空间,E 的点列(x_n)称为柯西序列,如果 $p,q\to\infty$ 时,$d(x_p,x_q)\to 0$,换言之,如果对于每个 $\varepsilon>0$,存在一整数 n_0,使得关系 $p,q\geqslant n_0$ 蕴涵 $d(x_p,x_q)<\varepsilon$.

这条件还可理解如下:用 X_n 表示集合$\{x_n,x_{n+1},\cdots\}$,如果 $n\to\infty$ 时,$\delta(X_n)\to 0$,则(x_n)是柯西序列.

显然,每一收敛序列都是柯西序列;一般说来,反之不真,例如,序列$\{1, 1.4, 1.41,\cdots\}$由十进制分数构成,它收敛于$\sqrt{2}$,在有理数集 $\mathbf{Q}$ 内是柯西序列,但在 $\mathbf{Q}$ 内它不收敛.

不过,确定一种尺度空间,使得柯西序列都收敛这一点是很重要的.由此有下述定义:

定义 5 如果尺度空间 E 中任何柯西序列都收敛,就说 E 是完备的.

刚才已知有理数空间 $\mathbf{Q}$ 不完备;反之,$\mathbf{R}$ 则是完备的,这在前面关于数直线的一讲中已经说过了.

R 中柯西序列的例子

1) 如果(a_n)下降于零,则交错级数 $a_1-a_2+a_3\cdots$ 的部分和 S_n 是柯西序列.

2) 如果级数 $|b_1|+|b_2|+\cdots$ 收敛,则级数 $b_1+b_2+\cdots$ 的部分和 S_n 是柯西序列.

完备空间与非完备空间的例子

1)$\mathbf{R}$ 的每个闭区间完备;$\mathbf{R}$ 的每个开集(除 $\mathbf{R}$ 本身外)非完备.

2) 欧氏空间 $\mathbf{R}^n$ 完备.事实上,$\mathbf{R}^n$ 的柯西序列在各坐标轴上的投影也是柯

西序列，因为投影的结果使距离缩短了．因此，这几个投影序列都收敛．从而所给柯西序列收敛．

3）每一紧致尺度空间完备．

完备性是一致同构的不变性

把 $\mathbf{R}$ 映成 $(-1,+1)$ 的映射 $x \to \frac{x}{1+|x|}$ 是一同胚，但是，它把完备尺度空间 $\mathbf{R}$ 映成非完备的尺度空间 $(-1,+1)$，所以完备空间的概念不是一个拓扑概念．其所以如此，是因为同胚并不总是把柯西序列变成柯西序列．

反之，把 E 映入 F 的每个一致连续映射一定把 E 的柯西序列变成柯西序列，所以两个空间 E 与 F 之间的一致同构映射使它们的柯西序列互相变换，特别，E 与 F 如果有一个完备，则另一个也完备．

因此，尺度空间的完备性在一致同构映射下是不变的；这是一致结构的一个性质．

下面是几个有用的性质，它们的证明留作练习：

1）在完备空间 E 中，直径趋于零的任何非空下降闭集序列 (F_n) 有非空交集，只有一个点组成，这点的每个邻域都包含序列中的一个集合．

2）完备空间的每个闭子集是完备空间．

3）尺度空间的每个完备子空间是该空间的闭集．

4）在尺度空间中，两个完备子空间的并集与交集是完备子空间．

一致连续函数的开拓

分析学常常提出一个问题：把定义在集 E 中处处稠密子集 A 上的函数连续地开拓到整个 E 上（处处稠密是指 E 的每个开集都含有 A 的点）．这种开拓的可能性乃是下列定理的推论：

定理 2　设 E 是尺度空间，A 是 E 的处处稠密子集，f 是把 A 映入完备尺度空间 F 的一致连续映射．于是，f 可以唯一地开拓成把 E 映入 F 的连续映射．开拓后的映射仍然一致连续（并且与 f 有一样的连续模）．

下面是证明提要：

对每个 $x \in E$ 与每个整数 n，令 $B\left(x, \frac{1}{n}\right)$ 是 E 中以 x 为心、$\frac{1}{n}$ 为半径的闭球．诸集合 $V_n = \overline{f(A \cap B(x, \frac{1}{n}))}$ 构成下降闭集序列，其直径趋于零（f 的一致连续性），所以存在唯一的公共点，记为 $g(x)$．



对于每个 $x \in A$，有 $g(x) = f(x)$. 另一方面，对每个 $X \subset E, g(x) \subset \overline{g(A \cap X)}$①，由此推出，$\delta(X) \to 0$ 时，$\delta(g(X)) \to 0$. 故 g 一致连续.

例子　1) 定义在有理数集 $\mathbf{Q}$ 上的 a^x 可以开拓至 $\mathbf{R}$.

2) 设 f 是定义在[0,1]上的可微数值函数，满足 $|f'| < k$，于是 f 是李普希兹函数，所以一致连续. 因此可以开拓成定义在[0,1]上的连续函数.

逐次逼近法

这里说的是在分析的许多领域(数值方程，微分方程，积分方程)中有广泛应用的一个有力的方法，它的长处先在E·比卡的工作中表现出来. 这方法可简要叙述如下：

首先，把尺度空间 E 映入另一个尺度空间 F 的映射 f 叫做 λ 压缩映射($0 \leqslant \lambda < 1$)，如果对 E 的任何两点 x_1, x_2，有

$$d_F(f(x_1), f(x_2)) \leqslant \lambda d_E(x_1, x_2)$$

这其实就是比值为 $\lambda < 1$ 的李普希兹映射.

例子　E 是欧氏空间 $\mathbf{R}^n$，f 是比值为 $\lambda < 1$ 的相似变换.

定理3　设 E 是完备尺度空间，f 是把 E 映入自身的压缩映射，则方程 $x = f(x)$ 有唯一解 a. 对于任何 $x_0 \in E$，点 x_0 的逐次变换序列 $x_n = f^n(x_0)$ 收敛于 a.

---

①　这个包含式显然不成立，例如 $X = E - A$ 时，$g(E-A) \not\subset \overline{g(A \cap (E-A))} = \overline{g(\phi)}$. 此外，本证明似乎没有这样简单，今据所给线索改证如下：

对于任何 $\varepsilon > 0$，据 f 的一致连续性，存在 $\delta > 0$，使得对任何 $a_1, a_2 \in A, d_E(a_1, a_2) < \delta$ 蕴涵 $d_F(f(a_1), f(a_2)) < \frac{\varepsilon}{6}$. 我们来证明：对任何 $X \subset E, \delta(X) < \frac{\delta}{3}$ 蕴涵 $\delta(g(x)) < \varepsilon$. 设 $x_1, x_2 \in X$，$d_E(x_1, x_2) < \frac{\delta}{3}$，按照 $g(x_1)$ 的定义，存在 $n_1 > \frac{3}{\delta}$，使得 $V_{n_1} \subset B(g(x_1), \frac{\varepsilon}{6})$ 对于任何 $y \in B\left(x, \frac{1}{n_1}\right)$，当 m 充分大时，$B\left(y, \frac{1}{m}\right) \subset B\left(x_1, \frac{1}{n_1}\right)$，所以

$$g(y) = \bigcap_m \overline{f\left(A \cap B\left(y, \frac{1}{m}\right)\right)} \subset \overline{f\left(A \cap B\left(x_1, \frac{1}{n_1}\right)\right)} = V_{n_1}$$

因此

$$g\left(B\left(x_1, \frac{1}{n_1}\right)\right) \subset B(g(x_1), \frac{\varepsilon}{6})$$

同理，存在 $n_2 > \frac{3}{\delta}$，使得

$$g\left(B\left(x_2, \frac{1}{n_2}\right)\right) \subset B(g(x_2), \frac{\varepsilon}{6})$$

于是

$$d_E(a_1, a_2) \leqslant d_E(a_1, x_1) + d_E(x_1, x_2) + d_E(x_2, a_2) < \delta$$

从而

$$d_F(g(x_1), g(x_2)) \leqslant d_F(g(x_1), f(a_1)) + d_F(f(a_1), f(a_2)) + d_F(f(a_2), g(x_2)) < \frac{\varepsilon}{2}$$

因此得到 $\delta(g(x)) < \varepsilon$. 最后，按照定义，$g(x) = \lim_{n \to \infty} f(a_n), a_n \in A \cap B\left(x, \frac{1}{n}\right)$，所以 $d_F(f(a_n), f(b_n)) \leqslant \varphi(d_E(a_n, b_n))$ 蕴涵 $d_F(g(x), g(y)) \leqslant \varphi(d_E(x, y))$，可见 g 和 f 有同样的连续模.

还应指出，上述证明中的记号 $B(x_1, \frac{1}{n_1})$ 等均遵照原文规定指闭球，但换成开球后，全部论证包括原文对 $g(x)$ 的定义也都仍然有效. 因此，原文规定 $B\left(x, \frac{1}{n}\right)$ 表示闭球似无必要.

证明 从关系 $x_{p+1}=f(x_p)$ 推出,对于每个 n,点对 (x_{n-1},x_n) 变成点对 (x_n,x_{n+1}).

因此 $d(x_n,x_{n+1})\leqslant\lambda d(x_{n-1},x_n)$.从前 n 个这样的关系得

$$d(x_n,x_{n+1})\leqslant\lambda^n d(x_0,x_1)$$

从而

$$d(x_n,x_{n+p})\leqslant d(x_0,x_1)\sum_n^{\infty}\lambda^p$$

因 $0\leqslant\lambda<1$,故序列 (x_n) 是柯西序列.

若 a 是 (x_n) 的极限点,因 f 连续,故有

$$f(a)=\lim f(x_n)$$

另一方面,$x_{n+1}=f(x_n)$,取极限得 $a=f(a)$.

这个 a 是方程 $x=f(x)$ 的唯一解,因为若 a' 是另一解,则有

$$d(a,a')=d(f(a),f(a'))\leqslant\lambda d(a,a')$$

由此 $d(a,a')=0$,故 $a=a'$.

注意,上述方法提供了解 a 的实际近似计算法,因为公项为 $d(x_n,x_{n+1})$ 的级数较之公比为 λ 的几何级数收敛得更快.确切地讲,我们有

$$d(x_n,a)=\lim_{p\to\infty}d(x_n,x_{n+p})\leqslant d(x_0,x_1)\frac{\lambda^n}{1-\lambda}$$

逼近序列的收敛还可能更快,例如,f 在 a 的一个邻域 V 上的限制是 $\lambda(V)$ 压缩映射,并且 V 的直径趋于0时 $\lambda(V)$ 趋于0的情形.

当空间 E 有界时,a 的存在与唯一性更明显,因为序列 $f^n(E)$ 下降,且直径 $\leqslant\lambda^n\delta(E)$.

例子 1) $E=\mathbf{R}^n$,f 是比值为 $\lambda<1$ 的相似变换,点 a 就是相似中心,而相似中心的存在就是这样证明的.

2) 设 $y=f(x)$ 是定义在有界闭区间 $E=[a,b]$ 上的可微数值函数,值域也含于 $[a,b]$ 内,并满足

$$|f'(x)|\leqslant\lambda\quad\lambda<1$$

把 $[a,b]$ 映入自身的映射 f 是 λ 压缩映射,逐次逼近法可以应用.

本文不可能来讨论这方法对微分方程的应用了.

尺度空间的完备化

上面对完备空间的讨论表明提出完备空间是有益的.当一种数学理论归结为考虑一个非完备的尺度空间时,我们几乎总是要在空间中添加新的元素,使之成为完备空间.理论上讲,完备化总是可能的,甚至在某种意下义,完备化还是唯一的.确切地讲,利用给定空间 E 的柯西序列,我们可以证明,存在一完备空间 $\hat{E}$,除同构不计外是唯一的,它包含一个与 E 同构的处处稠密子集(将这子





集与 E 看成一样是适宜的).

但是,如果我们不能指明添加到 E 的新元素,那么这种抽象的完备化一般没有什么用处;完备化的硬功夫就在于明确指出添加的新元素,有时这工作是困难的,可是常常很有益,因为这就表明理论发展的自然背景是什么.

例子 1) 设 E 是[0,1]上的数值连续函数所成的向量空间,它的距离由下列范数决定

$$\| f \| = \int_0^1 | f(t) | \mathrm{d}t$$

这个尺度空间不完备.相应完备空间 $\hat{E}$ 的元素是[0,1]上勒贝格可积函数(两个函数只在一零测度集合上不同时视为一样).

用下面的范数

$$\| f \|^2 = \int_0^1 f^2(t)\mathrm{d}t$$

代替上面的范数可以得到一个类似的例子.

这时新函数是平方可积的可测函数.

2) 设 E 是定义在 $\mathbf{R}^2$ 内,在一紧集之外为零的连续可微函数所成的向量空间,其范数定义为

$$\| f \|^2 = \iint \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right] \mathrm{d}x\mathrm{d}y$$

由此得到 E 上的距离.尺度空间 E 不完备;为完备化而添加到 E 的元素是具有若干有趣性质的新函数.

3 一致收敛的拓扑结构

迄今,我们都是研究一个函数的性态;进一步可以研究函数族的性态.这方面的研究开宗明义第一章就是函数序列逐点收敛的概念.

定义 6 设 f 和 $f_n(n = 1,2,\cdots)$ 是把集合 E 映入拓扑空间 F 的映射,如果对每个 $x \in E$,有 $f_n(x) \to f(x)$,就说序列 f_n 逐点收敛于 f.

虽然这种收敛方式在数学上曾经起过重大作用,现在也仍然有其一席之地,但是按照不同的需要还建立了另外一些收敛方式,至今仍然有用(见以下各例).

事实上,当我们讨论逐点收敛的函数序列的图象时,会发现这种收敛未能把我们对收敛的直观概念恰当地表达出来.

例如,设 f_n 是[0,1]上的数值函数,定义如下

$$f_n(x)=\begin{cases}n^2x(1-nx), & 当\ x\in\left[0,\frac{1}{n}\right]\\ 0, & 当\ x\in\left(\frac{1}{n},1\right]\end{cases}$$

序列 f_n 逐点收敛于函数 $f=0$，但 f_n 的图象却不收敛于 f 的图象，所以我们要引入一种新的收敛方法：

定义 7 设 f 和 $f_n(n=1,2,\cdots)$ 是把集合 E 映入尺度空间 F 的映射. 如果对每个 $s>0$，存在一整数 n_0，使得对于每个 $x\in E$，与每个 $n>n_0$，有 $d(f(x),f_n(x))<\varepsilon$，就说 f_n 一致收敛于 f.

在上面的例子中，f_n 不一致收敛于 0. 故逐点收敛不蕴涵一致收敛. 即使函数 f 与 f_n 连续，或者 E 是紧致集，或者 f_n 有界亦然.

反之，一致收敛显然蕴涵逐点收敛.

注意一个重要事实：上面的定义未假定定义域 E 有拓扑结构，但却要求值域 F 是尺度空间.

一致收敛的尺度结构与拓扑结构

设 $\mathfrak{F}(E,F)$ 是把 E 映入 F 的映射所成的集合（为简单计，假定空间 F 有界），我们可以在 $\mathfrak{F}(E,F)$ 中引入距离，使得对一致收敛的讨论更加直观、更加方便.

设 f 与 g 是把 E 映入 F 的两个映射. 如果对于每个 $x\in E$，有 $d(f(x),g(x))\leqslant\varepsilon$，就说 f 与 g 是 ε 相近；这些 ε 中的最小数称为 f 与 g 之间的距离，也就是

$$d(f,g)=\sup_{x\in E}d(f(x),g(x))$$

容易验明，$d(f,g)$ 的确是 $\mathfrak{F}(E,F)$ 上的距离；$\mathfrak{F}(E,F)$ 上这样定义的尺度称为一致收敛尺度.

现在可以对一致收敛给一个方便的解释：序列 (f_n) 一致收敛于 f 等价于 $d(f,f_n)\to 0$，换句话说，在 $\mathfrak{F}(E,F)$ 中的点列 (f_n) 收敛于点 f.

注意，如果把 F 的尺度换成另一个（关于一致结构）等价的尺度，那么在 $\mathfrak{F}(E,F)$ 上得到的两个尺度也是等价的，所以，一致收敛的概念只与 F 的一致结构有关.

连续函数的一致极限

如果我们现在假设定义域空间 E 具有拓扑结构，那么就可以研究一致收敛对连续函数的作用. 下面的基本定理的证明留作练习：

定理 4 设 E 是拓扑空间，F 是尺度空间. 如果把 E 映入 F 的连续映射序列一致收敛于映射 f，则 f 连续.

例子 设 (u_n) 是 $[0,1]$ 上的连续数值函数序列，对于每个 x 满足



$|u_n(x)| \leqslant a_n$,并且级数 $\sum a_n$ 收敛,则级数 $\sum u_n(x)$ 一致收敛(即部分和序列 $S_n(x)$ 一致收敛),其和是一连续函数.

级数 $\sum \frac{1}{3^n}\cos(2^n x)$ 即是一例.

注意,对于逐点收敛没有类似定理,例如,可以构造 $\mathbf{R}$ 上的连续函数序列,逐点收敛于一个函数 f: $x=0$ 时,$f(x)=0$; $x\neq 0$ 时,$f(x)=1$.

具有一致收敛尺度的连续函数所成的空间在数学上有重大的作用.特别,$[0,1]$ 上的连续数值函数空间 $C[0,1]$ 最常用;这是体 $\mathbf{R}$ 上的(无限维)向量空间,它的一致收敛尺度由范数 $\|f\| = \sup\limits_{x\in[0,1]} |f(x)|$ 产生.对于这个尺度,它是完备的;容易证明它不是局部紧致的.

4 结　论

本文中,我们碰到几个概念:函数的一致连续,柯西序列,完备空间,一致收敛.

这些概念都是从空间的尺度出发定义的;不过我们已经指出,将尺度换成一致同构的尺度后,它们是不变的,所以,这些概念不是尺度概念,可以期望推广到非尺度空间.

实际上,对于拓扑群或者更一般的所谓一致空间都可以定义一些类似的概念.拓扑群是指一个群,具有拓扑结构,使得映射 $x\to x^{-1}$,$(x,y)\to x\cdot y$ 连续.一致空间是指具有某种结构的集合,满足与尺度空间相应的几条公理,使我们可以界说所有子集的微小级.

第十一讲　某些函数空间与收敛方式的研究①

在前面有关拓扑空间的那些讲演中,我们是从一个实例(实数集)出发,从而上升到一般的公理.按照历史上的发展,还需要分析实数以外的一些例子,然后才能建立抽象的理论.其中,某些很重要的例子是函数空间提供的,所谓函数空间就是以函数作为元素的空间.函数空间在19世纪末就开始深入研究了.

本文中,我们只给出几个函数空间的例子,据以阐述前面几讲中提到的某些一般概念.我们不讲有关函数空间的任何一般定理.

1　第一个例子

用 C 表示定义在 $0 \leqslant x \leqslant 1$ 上的实值连续函数 $f(x)$ 的集合.若 $f \in C$, $g \in C$,我们令②

$$d(f,g) = \sup_{0 \leqslant x \leqslant 1} |f(x) - g(x)|$$

我们证明 d 是 C 上的距离.为此必须验明下列性质:

1) 对任意的 $f,g \in C, 0 \leqslant d(f,g) < \infty$.

2) $d(f,g) = 0$ 的充要条件是 $f = g$.

3) 对任意的 $f,g \in C, d(f,g) = d(g,f)$.

4) 对任意的 $f,g,h \in C, d(f,h) \leqslant d(f,g) + d(g,h)$.

性质1),2),3)是明显的.又设 $f,g,h \in C$,则对任意的 $x \in [0,1]$ 有

$$\begin{aligned}|f(x) - h(x)| &\leqslant |f(x) - g(x)| + |g(x) - h(x)| \leqslant \\ &\sup_{0 \leqslant x \leqslant 1} |f(x) - g(x)| + \sup_{0 \leqslant x \leqslant 1} |g(x) - h(x)| = \\ &d(f,g) + d(g,h)\end{aligned}$$

因此,$d(f,g) + d(g,h)$ 对任意 x 是 $|f(x) - h(x)|$ 的上界,所以是 $\sup_{0 \leqslant x \leqslant 1} |f(x) - h(x)| = d(f,h)$ 的上界.这就证明了性质4).

于是,就尺度空间讨论的所有概念都适用于空间 C.我们说,当 $n \to \infty$ 时,C 的函数序列 f_n 趋于 $f \in C$,表成 $d(f_n,f) \to 0$,意思是:对每个 $\varepsilon > 0$,存在一整数 N,使得关系 $n \geqslant N$ 蕴涵 $d(f_n,f) = \sup_{0 \leqslant x \leqslant 1} |f_n(x) - f(x)| \leqslant \varepsilon$.我们再考虑

① 原作者 J. Dixmier(索尔本大学讲师).
② 注意,$\varphi(x) = |f(x) - g(x)|$ 在一个闭区间上有上确界,用 $\sup \varphi(x)$ 表示.





一下 f_n 一致收敛于 f 的意义.

设(f_n)是 C 的函数序列.微分学中有一个经典定理,叫做柯西准则,给出了 f_n 一致收敛的充要条件如下:对每个 $\varepsilon > 0$,存在一整数 N,使得只要 $m \geqslant N$, $n \geqslant N$,则对任何 $x \in [0,1]$ 有 $| f_m(x) - f_n(x) | \leqslant \varepsilon$(换句话讲, $\sup\limits_{0\leqslant x\leqslant 1} | f_m(x) - f_n(x) | \leqslant \varepsilon$,亦即 $d(f_m,f_n) \leqslant \varepsilon$).于是$(f_n)$一致收敛的充要条件可说成:序列$(f_n)$是空间 C 中的柯西序列.另一方面,又一个经典定理告诉我们,连续函数序列的一致收敛极限是连续函数,由此可证 C 是完备空间.

定义在空间 C 上的实值函数是什么意思?这是一个函数 A,对于每个 $f \in C$,它对应一个实数 $A(f)$.函数 A 在 C 上连续意思是,当 f_n 一致收敛于 f 时,实数 $A(f_n)$ 在通常意义下趋于 $A(f)$.

例子　1) 令 $A(f) = \int_0^1 f(x)\mathrm{d}x$.根据一个经典的定理,$A$ 是 C 上的连续函数.2) 令 $A(f) = f(1)$.显然 A 是 C 上的连续函数.

现在考虑把 C 映入 C 的映射,即变换 $g = T(f)$:对于每个 $f \in C$,它对应一个 $g \in C$. T 连续的意思是:当 f_n 一致收敛于 f 时,$T(f_n)$ 一致收敛于 $T(f)$.例子:1) $T(f) = f_0 f$,这里 f_0 是一个固定的连续函数.显然,T 连续.2) $g = T(f)$ 定义如下

$$g(x) = \int_0^1 \mathrm{e}^{xy} f(y)\mathrm{d}y$$

g 是连续函数,我们证明 T 连续.设序列 $f_n \in C$ 一致收敛于 $f \in C$. $g_n = T(f_n)$, $g = T(f)$.我们有

$$\begin{aligned} | g_n(x) - g(x) | = & \left| \int_0^1 \mathrm{e}^{xy} f_n(y)\mathrm{d}y - \int_0^1 \mathrm{e}^{xy} f(y) \right| \mathrm{d}y \leqslant \\ & \int_0^1 \mathrm{e}^{xy} | f_n(y) - f(y) | \mathrm{d}y \leqslant \\ & \mathrm{e} \sup_{0\leqslant y\leqslant 1} | f_n(y) - f(y) | = \mathrm{e}d(f_n,f) \end{aligned}$$

所以,$d(g_n,g) \leqslant \mathrm{e}d(f_n,f)$,这就证明了 g_n 一致收敛于 g.

2　第二个例子

仍取上面的集合 C.对于 $f \in C, g \in C$,令

$$d'(f,g) = \int_0^1 | f(x) - g(x) | \mathrm{d}x$$

我们证明 d' 是 C 上的一个距离.上例中的性质 1),2),3) 是明显的.三角形不等式也易证:因为对每个 $x \in [0,1]$ 有

$$| f(x) - h(x) | \leqslant | f(x) - g(x) | + | g(x) - h(x) |$$

故

$$\int_0^1 |f(x)-h(x)|\mathrm{d}x \leqslant \int_0^1 |f(x)-g(x)|\mathrm{d}x + \int_0^1 |g(x)-h(x)|\mathrm{d}x$$

即
$$d(f,h) \leqslant d(f,g) + d(g,h)$$

这样,同一函数集合 C 具有第二个尺度空间结构.为区别计,我们用 C' 表示新得的尺度空间.在 C' 意义上的收敛称为平均收敛.若连续函数列 f_n 在 C 的意义下收敛于 f,即一致收敛于连续函数 f,则 f_n 在 C' 的意义下收敛于 f,因为 $|f_n(x)-f(x)|$ 一致趋于0,故 $\int_0^1 |f_n(x)-f(x)|\mathrm{d}x$ 趋于0.然而其逆不真,如下例所示:设 $f_n(x)=x^n$,$f(x)=0$,则

$$d'(f_n,f) = \int_0^1 x^n\mathrm{d}x = \frac{1}{n+1} \to 0 \quad (n \to +\infty)$$

故 f_n 在 C' 的意义下趋于 f;但是,f_n 不一致趋于 f,因为

$$\sup_{0\leqslant x\leqslant 1} |f_n(x)-f(x)| = \sup_{0\leqslant x\leqslant 1} x^n = 1$$

因此,在 C 上存在两种自然的收敛概念(其实有无限多,我们就不谈了).在函数空间中这种现象是很常见的.

现在证明 C' 不是完备的.定义连续函数 f_n 如下

$$f_n(x) = \begin{cases} 0, & \text{当 } 0 \leqslant x \leqslant \frac{1}{2}-\frac{1}{n} \\ 1, & \text{当 } \frac{1}{2}+\frac{1}{n} \leqslant x \leqslant 1 \end{cases}$$

当 $\frac{1}{2}-\frac{1}{n} \leqslant x < \frac{1}{2}+\frac{1}{n}$ 时,$f_n(x)$ 是线性函数.

设
$$f(x) = \begin{cases} 0, & \text{当 } 0 \leqslant x < \frac{1}{2} \\ \frac{1}{2}, & \text{当 } x = \frac{1}{2} \\ 1, & \text{当 } \frac{1}{2} < x \leqslant 1 \end{cases}$$

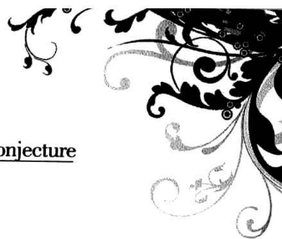
则
$$\int_0^1 |f_n(x)-f(x)|\mathrm{d}x = \frac{1}{2n} \to 0$$

故

$$\begin{aligned} d'(f_m,f_n) &= \int_0^1 |f_m(x)-f_n(x)|\mathrm{d}x \leqslant \\ &\int_0^1 |f_n(x)-f(x)|\mathrm{d}x + \int_0^1 |f(x)-f_m(x)|\mathrm{d}x = \\ &\frac{1}{2}\left(\frac{1}{n}+\frac{1}{m}\right) \to 0 \end{aligned}$$

所以,序列 (f_n) 是柯西序列.假定存在一个连续函数 g,使得 $d'(f_n,g) \to 0$,则





由三角形不等式得

$$\int_0^1 |f(x)-g(x)|\,dx \leqslant \int_0^1 |f(x)-f_n(x)|\,dx + \int_0^1 |f_n(x)-g(x)|\,dx$$

上述不等式右端,当 $n \to +\infty$ 时趋于0,故证得$\int_0^1 |f(x)-g(x)|\,dx = 0$.然而 $f(x)-g(x)$ 除去 $x=\frac{1}{2}$ 外连续,故 f 与 g 只在 $x=\frac{1}{2}$ 处不同.所以,$0 \leqslant x < \frac{1}{2}$ 时,$g(x)=0$;$\frac{1}{2} < x \leqslant 1$ 时,$g(x)=1$,这与 $g(x)$ 在 $x=\frac{1}{2}$ 处连续相矛盾.

(通过添加新函数,例如添加上面的 f,将 C' 扩大,便可得到完备空间.这是引入勒贝格意义下函数可积的手法)

再取第一个例子中考察过的函数 $A(f)=f(1)$.这函数在 C' 上不连续.事实上,设 $f_n(x)=x^n$,则在 C' 中 $f_n \to 0$. 然而,$A(f_n)=f_n(1)=1$ 不趋于 $A(0)=0$.反之,函数 $A(f)=\int_0^1 f(x)\,dx$ 在 C' 上连续.因为,如果 f_n 平均收敛于 f,则有

$$|A(f_n)-A(f)| = \left|\int_0^1 f_n(x)\,dx - \int_0^1 f(x)\,dx\right| \leqslant \int_0^1 |f_n(x)-f(x)|\,dx = d'(f_n,f) \to 0$$

明所欲证.

3 第三个例子

用 D 表示定义在 $0 \leqslant x \leqslant 1$ 上有连续导数的实值函数 $f(x)$ 的集合(对于 $x=0$ 是指右导数,对于 $x=1$ 是指左导数).若 $f \in D, g \in D$,令

$$\delta(f,g) = \sup_{0 \leqslant x \leqslant 1} |f(x)-g(x)| + \sup_{0 \leqslant x \leqslant 1} |f'(x)-g'(x)|$$

立即验明距离的头三条性质.三角形不等式由下列关系推出

$$\begin{aligned}
&|f(x)-h(x)| + |f'(x)-h'(x)| \leqslant \\
&|f(x)-g(x)| + |g(x)-h(x)| + |f'(x)-g'(x)| + \\
&|g'(x)-h'(x)| \leqslant \\
&\sup_{0 \leqslant x \leqslant 1} |f(x)-g(x)| + \sup_{0 \leqslant x \leqslant 1} |f'(x)-g'(x)| + \\
&\sup_{0 \leqslant x \leqslant 1} |g(x)-h(x)| + \sup_{0 \leqslant x \leqslant 1} |g'(x)-h'(x)| = \\
&\delta(f,g) + \delta(g,h)
\end{aligned}$$

所以 D 是尺度空间.

函数序列 $f_n \in D$ 收敛于函数 $f \in D$ 的意思是:f_n 一致收敛于 f,以及 f_n' 一致收敛于 f'(根据 $\delta(f_n,f)$ 的定义).

空间 D 是完备的. 因为,设(f_n) 是 D 中的柯西序列,则序列(f_n) 与(f_n') 满足柯西准则,故一致收敛于 f 与 g,一个经典定理断定 $g = f'$. 因此,f_n 在 D 的意义下收敛于 f.

对于每个 $f \in D$,令 $A(f) = f'(\frac{1}{2})$,则 A 是 D 上的连续函数.(但是,如果只是 D 的函数序列一致收敛于 D 的 f,则 $A(f_n)$ 不必收敛于 $A(f)$)D 上的这个连续函数 $A(f)$,就是许瓦兹在其《分布论》中所谓的双重泛函.

4 第四个例子

在复平面上,设 Δ 是由不等式 $|z| < 1$ 确定的开圆盘. 令 H 是函数 $f(z)$ 的集合,其中 $f(z)$ 是 Δ 上的全纯函数,满足 $\iint_{\Delta} |f(z)| \mathrm{d}\sigma < \infty$($\mathrm{d}\sigma$ 表示 Δ 内的面积元素). 对于 $f \in H$ 与 $g \in H$,令

$$d(f,g) = \iint_{\Delta} |f(z) - g(z)| \mathrm{d}\sigma$$

如同前面那些例子一样,可以验证 d 是距离(注意,因为 $\iint_{\Delta} |f| \mathrm{d}\sigma < \infty$, $\iint_{\Delta} |g| \mathrm{d}\sigma < \infty$,故 $d(f,g) < \infty$),所以 H 是尺度空间,H 中由此产生的收敛也称为平均收敛. 与第二个例子不同,下面要证明 H 是完备的.

首先证明,如果 $f \in H$,则对每个复数 ζ, $|\zeta| \leqslant a < 1$ 时,我们有

$$|f(\zeta)| \leqslant \frac{1}{\pi(1-a)^2} \iint |f(z)| \mathrm{d}\sigma \tag{1}$$

令 Γ_r 与 Δ_r 分别表示中心为 ζ,半径为 r 的圆周与开圆盘;对于 $r < 1 - a$,它们含于 Δ 内,令 $z = \zeta + r\mathrm{e}^{\mathrm{i}\theta}$ 便得

$$|2\mathrm{i}\pi f(\zeta)| = \left|\int_{\Gamma_r} \frac{f(z)}{z-\zeta} \mathrm{d}z\right| \leqslant \int_0^{2\pi} \frac{1}{r} |f(z)| r\mathrm{d}\theta = \int_0^{2\pi} |f(z)| \mathrm{d}\theta$$

由此得

$$\iint_{\Delta} |f(z)| \mathrm{d}\sigma \geqslant \iint_{\Delta_r} |f(z)| \mathrm{d}\sigma = \int_0^r \int_0^{2\pi} |f(z)| \rho \mathrm{d}\rho \mathrm{d}\theta \geqslant$$

$$\int_0^r 2\pi |f(\zeta)| \rho \mathrm{d}\rho = \pi r^2 |f(\zeta)|$$

由于上式对任意的 $r < 1 - a$ 成立,故得证(1).

现设(f_n) 是 H 中的柯西序列. 对每个满足 $0 \leqslant a < 1$ 的数 a,令 D_a 是中心为 0,半径为 a 的闭圆盘. 根据(1) 有





$$d(f_m,f_n) \geqslant \pi(1-a)^2 \mid f_m(\zeta) - f_n(\zeta) \mid \quad \zeta \in D_a$$

所以,f_n 在整个 D_a 内一致收敛于一函数f,根据魏尔斯特拉斯定理,此函数 f 在 Δ 内全纯.此外,对每个 $\varepsilon > 0$,存在正整数 N,对于 $m,n \geqslant N$ 以及任意的 $a < 1$ 有

$$\iint\limits_{D_a} \mid f_m(z) - f_n(z) \mid \mathrm{d}\sigma \leqslant \varepsilon$$

所以,对于 $n \geqslant N$ 以及任意的 $a < 1$ 有

$$\iint\limits_{D_a} \mid f(z) - f_n(z) \mid \mathrm{d}\sigma \leqslant \varepsilon$$

因此,对于 $n \geqslant N$,我们有

$$\iint\limits_{\Delta} \mid f(z) - f_n(z) \mid \mathrm{d}\sigma \leqslant \varepsilon$$

故 $d(f_n,f) \to 0$ 即 f_n 在空间 H 内趋于f.

5 第五个例子

设 B 是定义在区间$[0,1]$上的所有实值函数$f(x)$的集合. B 上最基本的收敛方式是逐点收敛:如果对每个 $x \in [0,1]$,$f(x)$ 趋于 $g(x)$ 就说f趋于g.我们将证明,在 B 上不可能定义出一个距离,使得相应的收敛是逐点收敛.这就表明,拓扑学不应该只研究尺度空间.

连续函数的集合 C 包含在B 中.设 C_1 是下述函数 $f \in B$ 的集合,这些f是 C 中函数序列逐点收敛极限. C_1 中的函数就是所谓的第一类函数.可以证明,这里我们权且承认:C_1 中的函数至少在一点连续①.要是逐点收敛可以由距离定义,则 C_1 就会是 B 中的闭集,即 C_1 是 C 在B 中的闭包.然而,我们却可以造出 C_1 的函数序列(f_n),逐点收敛于不在 C_1 中的一个函数f.作法分下列几步完成:

a) 函数$f \in B$ 除了在一点处等于 1 外,处处为 0,这样的f属于C_1(例如,f是一序列逐段线性连续函数的逐点收敛极限).

b) 函数$f \in B$ 除了在有限多个点处等于 1 外,处处为 0,这样的f属于C_1(f是有限多个 a) 型函数的和).

c) 设$(r_1,r_2,\cdots)$是排成一序列的有理数集②. f_n 是除了在点$r_1,r_2,\cdots,r_n$ 等于 1 外,处处为 0 的函数,则对每个 n,$f_n \in C_1$.

---

① 有了 C_1,就可以定义“第二类函数”,即是 C_1 中一函数序列的逐点收敛极限,这些函数的集合记为 C_2,…,一般可以对任何序数 α 定义 C_α.这是法国数学家贝尔(R.Baire)提出的函数分类概念.正是贝尔证明了:C_1 中任何函数的连续点构成$[0,1]$ 的处处稠密集.

② 因为在 0 与 1 之间的有理数集是可数的,所以可以排成序列.

d) 序列(f_n)逐点收敛于一个函数$f(x)$；x为无理数时，$f(x)=0$；x为有理数时，$f(x)=1$. 这个f没有任何连续点，所以不属于C_1.

在下面一讲中，我们将看到，可以在B上定义一个拓扑结构，使得相应的收敛就是逐点收敛.





第十二讲　一般拓扑的概念,拓扑空间的构造法①

为了研究极限与连续的概念,需要定义一个元素的邻近元素.现在,我们已经知道,有了距离 $d(x,x')$,就可以给出这样的定义.如果数列 $d(x_n,x)$ 趋于0,就说 x_n 趋于 x.因此,这一概念用到了实数集中极限的概念.然而,有一些性质与实数无关;另一方面,有些极限概念,即使很简单,却不能用距离来定义,所以就要求提出一个更一般的概念:元素的"邻域".

例如,考虑定义在区间 $[a,b]$ 中的函数 $u_i(x)$ 通常的收敛概念.如果对每个 $x\in[a,b]$,数 $u_i(x)$ 趋于0,就说 u_i 趋于0.如果对函数 $u_i(x)$ 的集合不加任何限制,这就表达了关于函数 u_i 的不可数无穷多个独立条件;容易证明,不可能找到定义在函数 u_i 的集合中的一个距离,使我们可以定义函数 u_i 趋于0.

在我们界定函数 $f(x)$ 在 x_0 处连续之前,已经知道,连续性可用下列方式定义:区间 $(f(x_0)-\eta,f(x_0)+\eta)$ 的逆象包含一个区间 $(x_0-\varepsilon,x_0+\varepsilon)$.任何集合,如包含一个含有 x_0 的开区间,就称为 x_0 的"邻域".因此,若 $f(x_0)$ 的任何邻域的逆像都是 x_0 的邻域,则 $f(x)$ 在 x_0 处连续.

因此,我们定义邻域时开集占有重要的地位.在尺度空间 E 的情形,开集 O 可以定义如下:如果对于 O 的每个元素 x_0,存在一数 $\rho>0$,使得所有满足 $d(x,x_0)<\rho$ 的 x 都属于 O,则 O 是开集.根据三角形不等式,任何集合 $\{x\mid d(x,x_0)<\rho\}$ 都是开集.这些开集满足下列三条与距离无关的性质.

O_1:任意(有限或无限多个)开集的并集是开集.

O_2:任意有限多个开集的交集是开集.

O_3:E 与 $\varnothing$ 是开集.

就 E 而言,O_3 是明显的;对于 $\varnothing$ 则是规定.对于 O_2,要注意"有限"一词是少不得的,这是因为,若考虑开集 $\{x\mid d(x,x_0)<\varepsilon_n\}$,这里 $\varepsilon_n\to0$,设它们有公共点 $x_1\neq x_0$,则有

$$d(x_1,x_0)<\varepsilon_n$$

故

$$d(x_1,x_0)=0,x_1=x_0$$

至于 O_1,几乎是明显的.

在任意集合 E 中,如果我们确定了 E 的一族子集,满足 O_1,O_2,O_3,就说在 E

① 原作者 Ch. Pisot (索尔本大学教授).

中定义了一个拓扑结构,或者说 E 成为一个拓扑空间.这些子集称为 E 的“开集”.

例子 考虑仅由 E 及 $\varnothing$ 构成的集族,它满足 O_1,O_2 与 O_3.相应的拓扑称为最粗的拓扑.

另一例子 设 E 是全序集合.所有满足 $x < a$ 或 $a < x < b$,或 $x > b$ 的元素 x 的集合称为“开区间”,开集就是有限多个或无限多个开区间的并集.

比如,考虑一个单调函数 $f(x)$.我们知道:对每个 x,存在右极限 $f(x+0)$ 与左极限 $f(x-0)$.于是,把 $\mathbf{R}$ 的每个实数 x 换成两个量,分别记作 $x+0$ 与 $x-0$,就从 $\mathbf{R}$ 得到一个集合 E,依下列序关系成为全序集:若在 $\mathbf{R}$ 内 $x < x'$,则在 E 内 $x+0 < x'-0$;又 $x-0 < x+0$.这就定义了 E 中的开集,从而定义了一个拓扑结构.

从开集的概念出发,可以定义邻域的概念.任何集合,如果包含一个含有元素 $x \in E$(或子集 $A \subset E$) 的开集时,就称为 x(或子集 A) 的邻域.

这概念表明开集是其每一点的邻域,反之亦然:任何集合 A,如果是它每一点的邻域,则 A 是开集.事实上,若 x 是 A 的一点,则存在一个开集 O_x,含有 x 且含于 A;A 显然是所有这些 O_x 的并集,故 A 是开集.

由此推出,如果知道了空间 E 中每一点的邻域,那么也就知道了 E 的所有开集.换言之,同一集合上的两个拓扑结构如果产生同样的邻域系,那么这两个拓扑结构是恒同的.

邻域具有下列性质,也可以作为出发点来定义拓扑结构:

V_1:x 的每一邻域含有 x.

V_2:每个包含 x 的一个邻域的集合是 x 的邻域.

V_3:x 的有限多个邻域的交集是 x 的邻域.

V_4:若 V 是 x 的一邻域,则存在 x 的邻域 W,使 W 含于 V,并且 V 是 W 的每一点的邻域.

性质 V_4 说明这样一个事实:任何一点,如果与 x 充分邻近的点充分邻近,它本身就与 x 邻近.这性质代替了尺度空间中的三角形不等式.为证明 V_4,只需取含于 V 且含有 x 的开集作 W 即可.

我们指出(但不详述),引入 E 的开集关于 E 的补集,可以得到一族集合,称为“闭集”,也可用来定义 E 的拓扑结构.

连续和极限 把拓扑空间 E 映入拓扑空间 F 的映射 f 称为在 E 的一点 u_0 处连续,如果对于 $f(u_0)$ 在 F 中的每个邻域 W.存在 u_0 在 E 中的邻域 V,使得 $f(V)$ 含于 W;或者说,$f(u_0)$ 在 F 中的每个邻域 W 的逆象 $f^{-1}(W)$ 是 u_0 在 E 中的邻域.如果映射 f 在 E 的每点连续,就说 f 在 E 上连续.f 在 E 上连续的充要条件是:F 的每个开集 O_F 的逆象 $f^{-1}(O_F)$ 是 E 的开集 O_E.事实上,O_F 是其各点的



邻域，因此，若f在E上连续，则$f^{-1}(O_F)$是其每一点在E中的邻域，所以是E中的开集. 反之，对E的每一点u以及$f(u)$的每一邻域W，存在开集O_F含有$f(u)$且含于W. 若对每个$O_F, f^{-1}(O_F) = O_E$，则u含于O_E，故$f^{-1}(W)$包含O_E，$f^{-1}(W)$是u的邻域，从而f连续.

我们说，序列u_n趋于u，如果对u的每一邻域V，存在一整数$N = N(V)$，使得$n \geqslant N$蕴涵u_n属于V. 遗憾的是一个序列可能不止一个极限，例如，按照最粗的拓扑结构，每个u都是任何序列的极限. 因此有必要特别提出下列拓扑空间：

隔离空间(或豪斯道夫空间)　E称为隔离的，如果它的邻域满足下列条件：E的两个不同的点具有两个不相交的邻域.

例子　E是全序集. 设$x < x'$是两个不同的元素，a与b使得$a < x < x' < b$.

(1) 存在c使得$x < c < x'$，于是(a, c)与(c, b)是两个不相交的开集.

(2) 不存在c使得$x < c < x'$，于是(a, x')与(x, b)不相交，因为，如果有一个公共的元素y，则$y < x', y > x$，故$x < y < x'$.

拓扑子空间　设E是拓扑空间，A是E的子集. E的开集与A的交满足O_1，O_2与O_3，所以得到A上的一个拓扑结构，使得把A映入E的包含映射连续. 包含映射就是对于任何$u \in A$相应地给出$u \in E$的映射.

例子　$\mathbf{R}$的区间(a, b)中的拓扑结构以及$\mathbf{R}^2$的子集圆周$x^2 + y^2 = 1$上的拓扑结构.

反过来讲，A的开集不一定是E中的开集. 要使A的开集都是E的开集，充要条件是：A是E的开集. 事实上，若A的开集都是E的开集，那么由于A是A的开集，故应是E的开集. 反之，若A是E的开集，则E的任何开集与A的交是E的开集，即是A的开集都是E的开集.

两个拓扑空间的积　设F与G是两个拓扑空间，$u \in F, v \in G$，于是$F \times G$就是所有点对(u, v)的集合. 考虑投影映射f与g，它们分别将$(u, v) \in F \times G$映成$u \in F$和$v \in G$. 我们在$F \times G$中引进一个拓扑结构，使得f和g连续. 于是对于F的开集$O_F, f^{-1}(O_F) = O_F \times G$应是$F \times G$中的开集；同样，对于$G$的开集$O_G, F \times O_G$也应是$F \times G$中的开集. 这些开集的有限交集与任意并集也应是开集. 如果我们将形如$O_F \times O_G$的积称为基本开集，则我们要找的开集就应是基本开集的任意并集，这些并集的确满足O_1, O_2与O_3. 事实上，O_1与O_3明显；今证O_2：设$A = \bigcup_\alpha O_F^\alpha \times O_G^\alpha, B = \bigcup_\beta O_F^\beta \times O_G^\beta$，由于$\cap$对$\cup$满足分配律，可见$A \cap B = \bigcup_{\alpha,\beta}(O_F^\alpha \times O_G^\alpha) \cap (O_F^\beta \times O_G^\beta) = \bigcup_{\alpha,\beta}(O_F^\alpha \cap O_F^\beta) \times (O_G^\alpha \cap O_G^\beta)$仍然是基本开集的并集. 这样得到的拓扑空间$F \times G$称为$F$与$G$的拓扑积. 我们还可依次定义若干个拓扑空间的积. 积是结合的.

连续性 设 φ 是把拓扑空间 E 映入拓扑积 $F\times G$ 的映射. 令 $f=\mathrm{proj}_F\varphi$ 与 $g=\mathrm{proj}_G\varphi$, 则 f 映 E 入 F, g 映 E 入 G. 若 φ 连续, f 与 g 也连续. 反之, 若 f 与 g 连续, 则 $f^{-1}(O_F)$ 和 $g^{-1}(O_G)$ 都是 E 的开集, 所以

$$\varphi^{-1}(O_F\times O_G)=f^{-1}(O_F)\cap g^{-1}(O_G)$$

是开集. 因此, $F\times G$ 中基本开集的任意并集, 即是它的任意开集在 φ^{-1} 下的象也都是 E 的开集, 故 φ 连续.

同样可知, $w_n=(u_n,v_n)\in F\times G$ 趋于 $w=(u,v)\in F\times G$ 的充要条件是: 在 F 中 $u_n\to u$, 在 G 中 $v_n\to v$.

例子 对于 $\mathbf{R}^2$, 开集是"方块"$(a,a')\times(b,b')$ 的并集.

若 F 与 G 是隔离空间, 则 $F\times G$ 亦然. 事实上, 若 (u,v) 和 (u',v') 是 $F\times G$ 中不同的点, 只要 u 与 u' 是 F 中不同的点, 则存在 $O_F\supset u$, $O'_F\supset u'$, 使 $O_F\cap O'_F=\varnothing$. 于是不论 G 中的点 v 与 v' 的邻域 O_G 和 O_G' 如何, 都有 $(O_F\times O_G)\cap(O'_F\times O'_G)=\varnothing$, 而 $O_F\times O_G$ 是 (u,v) 的邻域, 它与 (u',v') 的邻域 $(O'_F\times O'_G)$ 不相交.

反之, 容易证明, 若 $F\times G$ 是隔离空间, 则 F 与 G 亦然.

逐点收敛的空间 设 E 是数值函数 $u(x)$ 的空间, $x\in[a,b]$. 满足 $|u(x_i)-u_0(x_i)|<\varepsilon(i=1,2,\cdots,n)$ 的 $u\in E$ 的集合称为 E 中的基本开集, 记成

$$\omega=\omega(u_0,x_1,\cdots,x_n,\varepsilon)$$

有限多个或无限多个基本开集的并集称为"开集". 这样规定的开集满足性质 O_1, O_2 与 O_3 所得到的拓扑结构叫做逐点收敛拓扑结构.

这个拓扑结构是隔离的. 事实上, 若 $u_1\neq u_2$, 则至少存在一 x_0, 使得 $|u_1(x_0)-u_2(x_0)|=3\delta>0$, 于是 $w(u_1,x_0,\delta)$ 与 $w(u_2,x_0,\delta)$ 是两个不相交开集, 这是因为 $w(u_1,x_0,\delta)$ 是使得 $|u(x_0)-u_1(x_0)|<\delta$ 的 u 所成的集合.





第十三讲　紧致空间与局部紧致空间①

函数空间中极限的概念使我们特别重视具有下述性质的空间:其中任意无限多个函数至少有一极限函数.这个性质就是波尔查诺－魏尔斯特拉斯对有界无限实数集所说的那个性质.由于 M·蒙特尔的"正规族" 的理论,上述性质在复变函数论中的重要作用已经极为明显.可以指出,变分学对一般拓扑的建立也有巨大的贡献.

然而,对于函数空间提出的问题,波尔查诺－魏尔斯特拉斯定理的提法似乎并非总是最合适的形式.对于实数空间 $\mathbf{R}$,有一个与波尔查诺－魏尔斯特拉斯定理等价的、更令人满意的定理,就是波莱尔－勒贝格定理,这个定理才使我们能够定义一些极为有用的空间.

紧致空间　空间 E 称为紧致的,如果满足下列两个条件:

(1) E 是隔离的.

(2) 对于 E 的每个由开集组成的覆盖(开覆盖),可以从中选出有限多个开集,组成一个覆盖.

例如,$\mathbf{R}$ 的每一有界区间是紧致的②,但整个 $\mathbf{R}$ 并不如此.

紧致空间的定义可对偶地表述如下:E 称为紧致的,如果 E 是隔离的,并且对于 E 的任何一族闭集,只要其中任何有限子族都有非空的交,则整个集族也有非空的交,这后一定义显然表明,在紧致空间中,如果一族闭集 ϕ_i 按包含关系是全序族,则所有 ϕ_i 的交集非空,由此得到紧致空间 E 的波尔查诺－魏尔斯特拉斯定理.事实上,对于 E 的一族无限多个元素 a_n,令 $A_n=\{a_n,a_{n+1},\cdots\}$,$\bar{A}_n$ 是 A_n 的闭包,则 $\bar{A}_{n+1}\subset\bar{A}_n$.于是所有 $\bar{A}_n$ 的交集非空,这个交集的每个元素都是 a_n 所成集合的接触点.

这里要注意,逆命题一般不成立,即是具有波尔查诺－魏尔斯特拉斯定理所说性质的隔离空间不一定是紧致的.然而,后面会看到对于尺度空间逆命题是成立的.

容易看出,紧致空间的闭子空间是紧致的,同样,隔离空间的有限多个紧致子集的并集或交集是紧致空间.有限多个紧致空间的积(拓扑空间的积)是紧

---

① 原作者 Ch. Pisot(索尔本大学教授).

② 这里应该是:$\mathbf{R}$ 的任何有界闭区间是紧致的,例如 $\left(\frac{1}{n},1\right)$ $(n=2,3,\cdots)$ 显然是 $(0,1)$ 的开覆盖,但不可能选出有限子覆盖.

致空间;例如,n 维环面 T^n 是紧致的,因为圆周是紧致的.

连续性与紧致性　我们有下述定理:如果连续映射 f 把紧致空间 E 映入隔离空间 F,它就把 E 映成一个紧致空间 $f(E)$.

事实上,(1) 空间 $f(E)$ 是隔离的,因为它是隔离空间 F 的子空间.

(2) 设空间 $f(E)$ 有一个覆盖,由 $f(E)$ 的开集 ω_i 组成.这些开集可以表为 $\omega_i = f(E) \cap \Omega_i$,这里 Ω_i 是 F 的开集.由 f 的连续性推知,集合 $f^{-1}(\omega_i) = f^{-1}(\Omega_i)$ 是 E 的开集,它们构成 E 的覆盖.因 E 是紧致空间,故可选出有限子覆盖,经过映射 f,得到 $f(E)$ 的一个有限覆盖,由某些开集 ω_i 组成.故 $f(E)$ 是紧致空间.

特别,若 $F = \mathbf{R}$,即 f 是数值函数,则上述定理成立,所以一个紧致空间 E 在 $\mathbf{R}$ 中的连续象是紧致空间;特别,$f(E)$ 有界,并且 E 中存在某些元素,使 f 达到它的上确界与下确界.

紧致尺度空间　每个尺度空间都是隔离的,所以要使尺度空间是紧致空间,充要条件是波莱尔 - 勒贝格定理成立.我们已经知道,对于紧致尺度空间,这个条件蕴涵波尔查诺 - 魏尔斯特拉斯定理.对于尺度空间其逆亦真,亦即波尔查诺 - 魏尔斯特拉斯定理蕴涵紧致性.

事实上,考察尺度空间 E 的一个覆盖,由开集 ω_i 组成.对于 E 的每个元素 x 设 $\rho(x,i) \geqslant 0$ 是中心在 x 并含于 ω_i 的开球的最大半径,令 $\rho(x) = \sup \rho(x,i)$(就所有开集 ω_i 而言).若 $d(x,x')$ 表示 x 与 x' 之间的距离,则有

$$|\rho(x,i) - \rho(x',i)| \leqslant d(x,x')$$

故 $|\rho(x) - \rho(x')| \leqslant d(x,x')$;从而,$\rho(x)$ 是 E 上的连续函数.因为 E 中任何元素至少属于一个 ω_i,所以 $\rho(x) > 0$,因此波尔查诺 - 魏尔斯特拉斯定理蕴涵 $\rho(x)$ 在 E 上的下确界 $\rho = \inf \rho(x) > 0$.从而,对每个 x,存在指标 i_x 使得 $\rho(x,i_x) > \frac{\rho}{2}$.若 x_1 使得 $\omega_1 = \omega_{i_{x_1}}$ 不覆盖 E.我们取不属于 ω_1 的一个元素 x_2,如此继续下去①,得序列 $x_1, x_2, \cdots$ 没有接触点,这是因为序列中任何两个元素的距离都超过 $\frac{\rho}{2}$.于是,如果波尔查诺 - 魏尔斯特拉斯定理在 E 中成立,则上述序列是有限序列,从而从诸 ω_i 中选出了覆盖 E 的有限子族.

一致连续　设 E 是尺度空间,f 是把 E 映入尺度空间 F 的连续映射.若 E 是紧致空间,则映射 f"一致连续".事实上,对于任何 $x \in E$ 和任何 $\varepsilon > 0$,存在 $\eta_x > 0$,使得 $d_E(x,y) < \eta_x$ 蕴涵 $d_F(f(x),f(y)) < \varepsilon$.然而满足 $d_E(x,y) < \frac{1}{2}\eta_x$ 的点 y 的集合是 E 的开球 B_x,诸 B_x 构成 E 的覆盖.若 E 是紧致空间,便可

① 这里还应交代一句:若 $\omega_1 \cup \omega_2$ 不覆盖 E,则取 $x_3 \notin \omega_1 \cup \omega_2$.



从中选出有限覆盖$\{B_{x_i}\}$.

于是,令 $\eta = \min \frac{1}{2}\eta_{x_i}$;考虑 E 的一个元素 x,它属于某个 B_{x_i},例如 B_{x_1}. 设 y 是 E 的另一个元素,满足 $d_E(x,y) < \eta$,则

$$d_E(y,x_1) \leqslant d_E(y,x) + d_E(x,x_1) < \eta + \frac{1}{2}\eta_{x_1} \leqslant \eta_{x_1}$$

故 y 也属于 B_{x_1} 所以得

$$d_F(f(x),f(y)) \leqslant d_F(f(x),f(x_1)) + d_F(f(x_1),f(y)) < 2\varepsilon$$

连通性 如果对尺度空间 E 的任意两个元素 a 与 b,以及每个 $\varepsilon > 0$,可以相应求得 E 的有限多个元素 $a = a_0, a_1, \cdots, a_{n-1}, a_n = b$ 组成的链(可能与 ε 有关),使得 $d(a_i, a_{i+1}) < \varepsilon$ 就说 E 链式连通. 例如,**R** 的有界闭区间链式连通,**Q** 的有界闭区间亦然.

若 E 是链式连通的紧致尺度空间,f 是把 E 映成尺度空间 F 的连续映射,则 $f(E)$ 是紧致空间,f 一致连续,所以 $f(E)$ 链式连通.

一个链式连通的紧致尺度空间 E 是连通的,即 E 不能分成两个非空的不相交的闭集. 事实上,假若不然,设 E' 与 E'' 是这样两个非空的互不相交的闭集,则对 $x' \in E', x'' \in E''$ 我们有 $\inf d(x', x'') \neq 0$,否则存在 $x'_0 \in E'$ 与 $x''_0 \in E''$ 满足 $d(x'_0, x''_0) = 0$,故 $x'_0 = x''_0$,E' 与 E'' 就不是互不相交的了.

特别在 **R** 上,每个链式连通的紧致集是闭区间,所以,紧致尺度空间 E 上的数值连续函数取得 $\sup f(E)$ 与 $\inf f(E)$ 之间的一切值①.

局部紧致空间 空间 **R** 不是紧致的,然而它在每一点的邻近的性态却很像紧空间. 确切地讲,如果 E 是隔离的,并且 E 的每一点都具有紧致邻域,就说 E 是局部紧致空间. 例如,**R** 就是局部紧致的.

一个局部紧致空间可以添加唯一一个元素,即所谓无穷远点,而成为紧致空间. 为此,考虑 E 中的集合,其闭包是紧致集,这些集合的补集构成的集族 V 具有下列性质:任何集合,如果含有 V 的一个集合,本身就属于 V;V 中任意有限多个集合的交集属于 V. 两条性质就是用元素的邻域来定义拓扑结构时那个公理系统中的两条. 于是我们可以确定一个补充的元素 ω,规定族 V 是 ω 的邻域族. 这样,集合$\{E,\omega\}$就是紧致的了. 因为,设有$\{E,\omega\}$的一覆盖,考虑含有 ω 的开集 Ω 它的补集是紧致的,故可被有限多个开集 Ω_i 所覆盖,所以开集 Ω_i 与 Ω 就构成了$\{E,\omega\}$的有限覆盖②.

完备尺度空间 尺度空间叫做完备的,如何每一柯西序列都收敛. 实数集 **R** 是完备的,但有理数集 **Q** 不是完备的. 我们有下述命题:

---

① 空间 E 还应是链式连通的.

② 还应指出:$\{E,\omega\}$ 是隔离的,只需考虑 $x \in E$ 和 ω,这是显然的,因为 E 局部紧致;此外,E 作为$\{E,\omega\}$的子空间是处处稠密的,而且同胚于原来的空间 E. 这样得到的$\{E,\omega\}$称为 E 的单点紧致扩充.

任何紧致尺度空间都是完备的.

事实上,设(a_n)为柯西序列,即对每一$\varepsilon>0$,存在n使得$h\geqslant n,k\geqslant n$时,$d(a_h,a_k)<\varepsilon$.令$A_m=(a_m,a_{m+1},\cdots)$,$\bar{A}_m$是$A_m$的闭包,则$\bar{A}_{m+1}\subset\bar{A}_m$.若$a'$与$a''$是$\bar{A}_m$的两个元素,则有

$$d(a',a'')\leqslant d(a',a_h)+d(a_h,a_k)+d(a_k,a'')\quad h\geqslant n,k\geqslant n$$

又对任何$\varepsilon'>0$,存在$h\geqslant n,k\geqslant n$,使得$d(a',a_h)<\varepsilon',d(a_k,a'')<\varepsilon'$.因$(a_n)$是柯西序列,故

$$d(a',a'')<2\varepsilon'+d(a_h,a_k)<\varepsilon$$

命A是所有$\bar{A}_m$的交集.因空间紧致,故A非空.设a与b是A的两个元素,它们属于每个$\bar{A}_m$,故对任意$\varepsilon>0$有$d(a,b)<\varepsilon$,所以$d(a,b)=0,a=b$,A由唯一元素a组成.故柯西序列有唯一接触点,从而收敛.

巴拿赫空间　下面将定义体K中绝对值的概念.设E是体K上的向量空间,K中定义了$\alpha\in K$的绝对值,用$|\alpha|$表示.把E映入正实数集的映射$x\to\|x\|$称为范数,如果满足下列条件:

(1)$x\neq 0$时,$\|x\|>0$;$\|0\|=0$.

(2)对每一$\alpha\in K$有$\|\alpha x\|=|\alpha|\,\|x\|$.

(3)$\|x+y\|\leqslant\|x\|+\|y\|$.

这时,空间E称为赋范向量空间.

如果在这样的空间中,令

$$d(x,y)=\|x-y\|$$

它就成为尺度空间.

一个体K总可视为自身上的向量空间,如果它是赋范向量空间,则元素$\alpha\in K$的范数用$|\alpha|$表示,称为α的绝对值.

定义　在具有绝对值的体上的赋范向量空间,如果是完备尺度空间,就称为巴拿赫空间.

赋范向量空间的完备化　每个不完备的赋范向量空间E总可扩张成巴拿赫空间$\hat{E}$.

事实上,令$X=(x_1,\cdots,x_n,\cdots)$是$E$中的柯西序列,即对任何$\varepsilon>0$,存在一个$n$,使得$h\geqslant n,k\geqslant n$时$\|x_h-x_k\|<\varepsilon$.令$Y=(y_1,\cdots,y_n,\cdots)$是另一柯西序列,令

$$\alpha X=(\alpha x_1,\cdots,\alpha x_n,\cdots)\quad\alpha\in K$$

$$X+Y=(x_1+y_1,\cdots,x_n+y_n,\cdots)$$

易知αX与$X+Y$仍是柯西序列,故所有柯西序列构成K上一个新的向量空间.

我们有

$$\|x_h\|=\|x_h-x_k+x_k\|\leqslant\|x_h-x_k\|+\|x_k\|$$



故 $h \geqslant n, k \geqslant n$ 时

$$\|x_h\| - \|x_k\| \leqslant \|x_h - x_k\| < \varepsilon$$

交换 h 与 k,同样有

$$\|x_k\| - \|x_h\| \leqslant \|x_k - x_h\| < \varepsilon$$

因此实数序列 $\|x_1\|, \cdots, \|x_n\|, \cdots$ 是柯西序列,即它收敛于一实数 $\geqslant 0$.

现在令 $\|x\| = \lim_{n \to \infty} \|x_n\|$,则有

$$\|\alpha x\| = \lim_{n \to \infty} \|\alpha x_n\| = |\alpha| \|X\|$$

$$\|X + Y\| = \lim_{n \to \infty} \|x_n + y_n\| \leqslant \|X\| + \|Y\|$$

但是,$\|X\| = 0$ 并不蕴涵 $X = 0 = (0, \cdots, 0, \cdots)$,故 $\|X\|$ 不是范数.事实上,考虑柯西序列 $Z = (z_1, \cdots, z_n, \cdots)$,这里 $\lim_{n \to \infty} \|z_n\| = 0$,令 ζ 是所有这些柯西序列的集合,所以其中每个序列都满足 $\|Z\| = 0$.若 $Z \in \zeta, Z' \in \zeta$,则 $\alpha Z \in \zeta, Z + Z' \in \zeta$;关系 $X' \sim X \Leftrightarrow X' - X \in \zeta$ 是一等价关系.设 ξ 是 X 的等价类,η 是 Y 的等价类,令 $\alpha\xi$ 是 αX 的等价类,$\xi + \eta$ 是 $X + Y$ 的等价类;这些运算对等价类有定义,从而所有这等价类的集合构成一向量空间 $\hat{E}$.等价类 ζ 是 $\hat{E}$ 的零元素.

若 $X' \sim X$,则有 $X' = X + Z, Z \in \zeta$,故 $\|X'\| \leqslant \|X\|$(因 $\|Z\| = 0$),同样 $\|X\| \leqslant \|X'\|$,所以 $\|X''\| = \|X\|$.令 $\|\xi\| = \|X'\|$,则 $\|\xi\|$ 是范数;事实上,性质(2)和(3)已经证明;(1)也成立,因为若 $\|\xi\| = 0$,则 $\|X\| = 0$,故 $\lim_{n \to \infty} \|x_n\| = 0, x \in \zeta$,所以 $\xi = \zeta$.

若 $x \in E$,我们令 $\hat{x}$ 是$(x, \cdots, x, \cdots)$的等价类,于是,$\widehat{\alpha x} = \alpha\hat{x}$,$\widehat{x + y} = \hat{x} + \hat{y}$ 故 E 同构于 $\hat{E}$ 的一个子空间,将 $\hat{E}$ 的这个子空间与 E 等同,所以 $\hat{E}$ 是包含 E 的赋范向量空间.

将上面的过程用到体 K 得到向量空间 $\hat{K}$,可以证明:$\hat{K}$ 本身是一个体.若$(\alpha_1, \cdots, \alpha_n, \cdots)$是 K 中元素的绝对值构成的柯西序列,其等价类 α 属于$\hat{K}$.另一方面,序列$(\alpha_1 x, \cdots, \alpha_n x, \cdots)$是 E 的柯西序列,故它的等价类属于 $\hat{E}$,用 $a\hat{x}$ 表示.由此不难推出 $\hat{E}$ 是$\hat{K}$ 上的向量空间.

向量空间 $\hat{E}$ 是完备的 事实上,令$(\xi_1, \cdots, \xi_n, \cdots)$是 $\hat{E}$ 的柯西序列,即对每一 $\varepsilon > 0$,存在 n,使得

$$h \geqslant n, k \geqslant n \text{ 时}, \|\xi_h - \xi_k\| < \varepsilon$$

令 ξ_m 是 X_m 的等价类,这里 $X_m = (x_{m,1}, \cdots, x_{m,p}, \cdots)$,则 $p \geqslant N_m$ 时

$$\|\xi_m - \hat{x}_{m,p}\| = \lim_{q \to \infty} \|x_{m,q} - x_{m,p}\| < \varepsilon$$

相应于每个 ξ_m,可以得到元素 $\hat{y}_m$,满足 $y_m = x_{m,N_m}$,于是 $\|\xi_m - \hat{y}_m\| \leqslant \varepsilon$,所以:

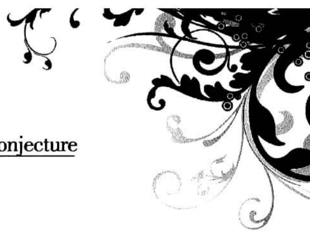
对于 $h \geqslant n, k \geqslant n$, $\| \hat{y}_h - \hat{y}_k \| \leqslant \| \hat{y}_h - \xi_h \| + \| \xi_h - \xi_k \| + \| \xi_k - \hat{y}_k \| < 3\varepsilon$. 而 $\| \hat{y}_h - \hat{y}_k \| = \| y_h - y_k \|$,故序列 $Y = (y_1, \cdots, y_n, \cdots)$ 是 E 的柯西序列. 若 η 是 Y 的等价类,则有

$$\| \eta - \xi_m \| \leqslant \| \eta - \hat{y}_m \| + \| \hat{y}_m - \xi_m \|$$

但是 $m \geqslant n$ 时, $\| \eta - \hat{y}_m \| = \lim\limits_{p \to \infty} \| y_p - y_m \| \leqslant 3\varepsilon$,而 $\| \hat{y}_m - \xi_m \| \leqslant \varepsilon$,所以 $m \geqslant n$ 时, $\| \eta - \xi_m \| \leqslant 4\varepsilon$. 即 $\xi_m \to \eta$,空间 $\hat{E}$ 完备.

应用 考虑有理数体 $\mathbf{Q}$. 研究 $\mathbf{Q}$ 上所有可能的绝对值,我们发现:除了通常的绝对值以外,还有下面的绝对值:设 $p \geqslant 2$ 是一固定素数,令 $r = \frac{u}{v}$ 是一有理数. 可以把$\frac{u}{v}$ 写成形式 $p^n \frac{u'}{v'}$,这里 u', v' 不能被 p 整数,指数 n 是正,负整数或零. 令 $| r |_p = \frac{1}{p^n}$ 容易验证这是绝对值. 按照通常的绝对值使 $\mathbf{Q}$ 完备化便得 $\mathbf{R}$. 按照绝对值 $| r |_p$ 使 $\mathbf{Q}$ 完备化则得到一个新的完备体 $\mathbf{R}_p$,在 $\mathbf{R}_p$ 上我们可以构造一种分析,使我们有可能得到实数分析得不出的结果.





第十四讲　代数结构与拓扑结构的相容性·拓扑群与拓扑向量空间[①]

标题所示内容很广泛，例如，包括拓扑群，函数空间等．我们将只讨论拓扑向量空间的几个重要性质，略见一斑．

1　拓扑向量空间的定义

复习一下定义在一个体(我们只用实数体$\mathbf{R}$,复数体的情形类似)上的向量空间E:E关于一个交换运算构成群，这个运算记为加法(对于E的一对元素x,y,相应地得到E的元素$x+y$,称为这两个元素的和)；E上还定义有E的元素与$\mathbf{R}$的元素的乘法：对于$x \in E$与$\lambda \in \mathbf{R}$,相应地有$\lambda x \in E$,这个运算法则对于$\mathbf{R}$的元素是结合的，对于$\mathbf{R}$中的加法与E中的加法是分配的，并且$1 \cdot x = x$.

E的元素x称为点，或者向量，在几何表示中，如果把一个向量的起点放在O,则该向量同它的终点将不予区别，如同复数的表示那样．

再假定E上还定义了一个拓扑结构．就是说，在E上确定了一族开集，或一族闭集，或者点的邻域，或者说明了序列的收敛性条件或函数的连续性条件．

但是，要得到有意义的特性，必须要有某些关系，使原有的代数结构与引入的拓扑结构相容：拓扑向量空间就是具有拓扑结构的向量空间，满足下列条件：

1) 函数$x+y$对x与y连续．

2) 与3) 关于$\mathbf{R}$上通常的拓扑结构和E上引进的拓扑结构，函数λx对λ与x连续．

这些条件可用序列的收敛性条件表成：若$x_1, x_2, \cdots, x_n, \cdots$收敛于$x$,$y_1, y_2, \cdots, y_n, \cdots$收敛于$y$,则序列$\{x_n + y_n\}$应收敛于$x+y$．还有，若序列$\lambda_1, \lambda_2, \cdots, \lambda_n, \cdots$收敛于$\lambda$,则序列$\{\lambda_n x_n\}$应收敛于$\lambda x$.

也可以从E的开集出发表达上述定义中的条件：我们要侧重讨论的就是这方面的内容．

① 原作者 R. Godement (索尔本大学教授).

2 用开集条件定义拓扑向量空间的公理

设给定 $a \in E$,对于 E 的每个元素 x,可以相应地得到 $x + a$(由 x 平移而得).若 x 是 E 的子集 U 的一般元素,我们用 $U + a$ 表示点 $x + a$ 所成的集合.映射 $x \to x + a$(及其逆 $x \to x - a$)应连续,所以,U 是开集等价于对任意 a,$U + a$ 是开集.

同样,对于 $\lambda \neq 0$,映射 $x \to \lambda x$ 应该对 x 连续,特别,若 x 是 O 的一个邻域内的一般点,λx 也是一个邻域内的一般点.

由此提出下列公理:

公理 1 关于任意向量的平移,E 的开集变成开集.关于中心在 O、比值任意的位移,O 的邻域变成 O 的邻域.

因此,为了确定 E 上的拓扑结构,只需说明原点 O 的所有邻域的特性,以保证条件 1),2) 与 3) 在原点成立.条件 3) 已经说明,还剩下条件 1) 和 2).

函数 $x + y$ 在原点连续,意思是:对 O 的任意邻域 U,存在 O 的两个邻域 V 与 W,使得

$$x \in V, y \in W \Rightarrow x + y \in U$$

亦即 $V + W \subset U$.

现设 U' 是 V 与 W 的交集,这是 O 的邻域,故上面条件等价于:

公理 2 对 O 的任意邻域 U,存在 O 的邻域 U',使得 $U' + U' \subset U$.

最后还要说明映射 $\lambda \to \lambda x_0$ 与 $(\lambda, x) \to \lambda x$(在点 $\lambda = 0, x = 0$)的连续性.

为了简化表达,一个集合 X 如果与任何一条通过 O 的直线的交集是中心在 O 的一个非空区间,就说 X 是星状对称集.利用这一术语,最后的两个条件可表述成:

公理 3 在 O 的每个邻域内,存在一个星状对称的子邻域.

换句话讲,O 的所有星状对称邻域构成 O 的一个邻域基.

3 局部凸的拓扑向量空间

在平面内,按照通常的拓扑,一个集合 A 叫做凸集,如果对于 A 的任意两点 a 与 b,线段 ab 上任意一点仍属于 A,同样,根据直线段的分析定义,对于实数体上的任何向量空间,子集 A 叫做凸集,如果

$$\forall x, y \in A, \forall \lambda \in [0,1]$$

我们有

$$\lambda x + (1 - \lambda) y \in A$$

一个拓扑向量空间叫做局部凸的,如果原点的每一邻域都包含原点的一个





凸邻域.这个条件在平移与位似作用下保持不变.

带有通常尺度的空间 $\mathbf{R}^n$ 显然满足这个条件,因为 O 的每个邻域包含一个中心在 O 的球,球当然是凸邻域.实际上,已经遇到过的那些拓扑向量空间几乎都是局部凸的,特别在分析中是这样.在代数中倒未必如此.

4 赋范空间

实数体上一个拓扑向量空间中定义的半范数是一个数值函数 $p(x)$,满足下列条件:

1) $p(x)$ 是一正数或零;

2) $p(x+y)\leqslant p(x)+p(y)$;

3) $p(\lambda x)=|\lambda|p(x)$.

特别,注意 $p(-x)=p(x)$, $p(0)=0$.

此外,若 $p(x)$ 仅当 $x=0$ 时为 0,即

4) $p(x)=0\Rightarrow x=0$,则函数 $p(x)$ 叫做范数,记成

$$p(x)=\|x\|$$

具有范数的向量空间叫做赋范空间.

向量空间上如果存在范数,则可以导出相应的拓扑结构,两个点 x 与 y 之差的范数称为两点的距离:$d(x,y)=\|x-y\|$ 这个数满足距离的公理,从而这个向量空间成为尺度空间.于是可定义球是满足 $d(x,a)<\rho$ 的 x 的集合,它可看做开集,而 a 的邻域定义成一个子集,它包含一个含有 a 的球.

剩下要验明,这样定义的拓扑结构与向量空间的结构是相容的

$$\begin{aligned}d(x+y,x_0+y_0)=&\|(x+y)-(x_0+y_0)\|=\\&\|(x-x_0)+(y-y_0)\|\leqslant\\&\|x-x_0\|+\|y-y_0\|\end{aligned}$$

即是

$$d(x+y,x_0+y_0)\leqslant d(x,x_0)+d(y,y_0)$$

而

$$\left.\begin{matrix}x\to x_0\\y\to y_0\end{matrix}\right\}\Leftrightarrow\left\{\begin{matrix}d(x,x_0)\to 0\\d(y,y_0)\to 0\end{matrix}\right\}\Leftrightarrow d(x+y,x_0+y_0)\to 0$$

故

$$\left.\begin{matrix}x\to x_0\\y\to y_0\end{matrix}\right\}\quad\text{蕴涵}\quad x+y\to x_0+y_0$$

同样,$x\to x_0$ 与 $\lambda\to\lambda_0$ 蕴涵 $\lambda x\to\lambda_0x_0$.

5 赋范空间的例子

显然,空间 $\mathbf{R}^n$ 是赋范空间.但是,这个概念对于函数空间尤其重要,函数空

间是无穷维的,元数可以是函数、测度等.

例 1 命 E 是实变数 t 的有界实函数 $x(t)$ 的集合,例如对 $t \in [0,1]$ 有定义.对于每个这样的函数 x,其绝对值的上确界是确定的,所以是 x 的函数

$$K(x) = \sup_t |x(t)| < +\infty$$

显然,$K(x)$ 是正的,是函数 x 所成集合中的范数.事实上

$$K(x+y) = \sup_t |x(t)+y(t)| \leqslant \sup_t(|x(t)|+|y(t)|) \leqslant K(x)+K(y)$$

$$K(\lambda x) = \sup_t(|\lambda||x(t)|) = |\lambda| K(x)$$

最后,$K(x) = 0$ 蕴涵对任意 t 有 $|x(t)| = 0$.

如上定义的范数,我们已经说过,在 E 上引入一相应的拓扑结构;序列 $x_n(t)$ 的收敛由条件 $d(x,x_n) \to 0$ 确定.这里

$$d(x,x_n) = \| x - x_n \| = \sup_t |x(t) - x_n(t)|$$

一个函数的上界小于 ε,就是说,对于任意 t,函数小于 ε,于是上述收敛性可以表成:

$\forall \varepsilon > 0, \exists N$,使得 $n > N$ 蕴涵对每个 $t \in [0,1]$ 有

$$|x(t) - x_n(t)| < \varepsilon$$

这是我们熟知的一致收敛的特性.

另一些例子 可以考虑上述空间的某些子集合,例如[0,1]上的连续函数的集合(这种函数是有界的),可微函数的集合,n 次可微函数的集合,或者无限可微函数的集合.

也可以改变范数的定义,例如,[0,1]上的连续函数集合中,可以取

$$\| x(t) \| = \int_0^1 |x(t)| \mathrm{d}t$$

它自然满足范数公理.

6 局部凸拓扑向量空间的结构

我们来证明,这样的空间不总是赋范空间.为了得到最一般的空间,必须从一族半范数 $p_\alpha(x)$ 出发,诸 α 是参数,可以有无限多个.

从这族半范数中,选出有限多个半范数 $p_{\alpha_1}, p_{\alpha_2}, \cdots, p_{\alpha_n}$,相应于每个半范数 p_{α_i} 给出一个正数 ε_i,然后,得到原点的一个邻域,即是满足 $p_{\alpha_i}(x) < \varepsilon_i$ 的那些 x 的集合,记为 $B(p_{\alpha_i}, \varepsilon_i)(i = 1,2,\cdots,n)$.最后,令 A 是这 n 个邻域的交集.

因此,对于任何有限多个参数 α,都可以相应地得到原点的一个邻域 A.我们将这些 A 所成的集合 $\mathscr{A}$ 取作 O 的基本邻域系(即 E 的每一子集,如包含一个



集合 A,就是 O 的邻域).这就把只有一个范数的情形推广了.

现证这样确定的拓扑结构使 E 成为局部凸空间.为此只需证明诸 A 是凸集,这只要诸 $B(p_{\alpha_i},\varepsilon_i)$ 是凸集就行了.因为凸集的交集是凸集.今设 $p_{\alpha_i}(x)<\varepsilon_i$,$p_{\alpha_i}(y)<\varepsilon_i$,对于 $0\leqslant\lambda\leqslant 1$,根据半范数的公理有

$$p_{\alpha_i}[\lambda x+(1-\lambda)y]\leqslant p_{\alpha_i}(\lambda x)+p_{\alpha_i}[(1-\lambda)y]\leqslant$$
$$|\lambda|p_{\alpha_i}(x)+|1-\lambda|p_{\alpha_i}(y)<$$
$$\lambda\varepsilon_i+(1-\lambda)\varepsilon_i=\varepsilon_i$$

反之,可以证明,任何局部凸的拓扑向量空间 E 上,存在一族半范数,可以按照刚才所讲的方法重新得到 E 上原有的拓扑结构.

例子 单个实变数的无穷可微函数 $x(t)$.(这个讨论可以推广到多变量函数)为了定义其分布,L·许瓦兹还假定这些函数当 t 超过某一有限值(与特定函数有关)时为零.我们这里不作这一补充假设.

显然,这些函数的集合 E 是向量空间,还可以赋予集 E 某些拓扑结构,例如,可以定义,如果对每一 t 值,$x_n(t)$ 收敛于 $x(t)$,就说 x_n 收敛于 x.这样的定义没有用到 $x(t)$ 可微的假设,下面引入一种很强的拓扑结构:$x_n(t)$ 叫做收敛于 $x(t)$,如果:1)在任何有限区间$[t_1,t_2]$上,$x_n(t)$一致收敛于 $x(t)$.还有2)各阶导数满足同样的条件:在长度有限的任何闭区间上,对任意正整数 h,$x_n^{(h)}(t)$ 一致收敛于 $x^{(h)}(t)$(注意,导数的收敛蕴涵函数的收敛,反之未必).

利用这种收敛的定义,可以引入一族无限多个半范数,从而在 E 上确定一拓扑结构.

令 K 是有限长的区间$[t_1,t_2]$,即 $\mathbf{R}$ 的一个紧致集.用 D 表示随便哪一阶导数的符号,D 是一族算子

$$x\to x^{(h)}=Dx,\forall h$$

我们定义关于 K 与 D 的半范数如下

$$p_{K,D}(x)=\sup_{t\in K}|Dx(t)|(|Dx|\text{ 在 }K\text{ 上的最大值})$$

这里,参数 α 是一对参数K,D.这样定义的半范数的确满足半范数公理,而由此产生的收敛也满足所要的强收敛条件,因为:

$\forall K,\forall D,p_{K,D}(x_n-x_0)$ 趋于 $0\Leftrightarrow\forall h,x_n^{(h)}(t)$ 在 K 上一致趋于 0.

还可以给出别的例子.

7 局部凸的拓扑向量空间的重要性

历史梗概 数学理论之所以得到发展,是因为可以用来研究某些精密的科学问题.我们刚才介绍的理论中有些定理威力极大,但为数不多.这就说明这

些定理的发现何以姗姗来迟(1920年左右).

巴拿赫开头只考虑一个范数;巴拿赫空间就是指完备赋范空间(这就是说,这个空间中所有的柯西序列都收敛).1930年左右所知道的那些重要定理,虽然只有五六个,但却适用于许许多多不同的情况.不过25年,人们终于能够选出那些好的公理,即是在那些有用的例子里成立的公理,并且把适用于这些例子的定理收集起来.巴拿赫的著作(《Théorie des Opérations Linéaires》,Varsovie,1932)标志着一个阶段的开始,只是在十五年以后才告终了,那是为了需要使这一理论能够兼收并蓄,研究更一般的集合,例如必须包容具有无限多个半范数的无限可微的许瓦兹空间.J·迪厄多内的一篇文章(见 Annales de l'E. N. S.,1940)重新考虑了巴拿赫的问题,引进了空间是局部凸的假设,从而可以得到一些定理,即使就巴拿赫研究过的情形而言,这些定理也比巴拿赫得到的结果要广(这里指G. W. Mackey, J. Dieudonné, L. Schwartz关于分布理论的工作).最后,Grothendieck应该看成是巴拿赫的真正接班人,因为他得到的那些结果非常重要.

有一些理论如果不从应用着眼就无法建立,也不可能发扬光大,局部凸空间的理论可以称得上是其中的典型.

8　简要介绍几个定理

汉恩 - 巴拿赫定理　这个定理引人注目,就连不是这方面的专家也深信这一理论的重要.我们不介绍它的证明了,这要用到超限归纳法与策墨诺公理.

准备　把一个拓扑向量空间映成另一个拓扑向量空间的连续线性映射.

把一个拓扑向量空间E映成另一个拓扑向量空间F的线性映射中,我们考虑关于E与F的拓扑结构是连续的那些映射.这类映射在许多场合是极重要的,例如在许瓦兹空间中,由导数定义的映射$x(t)\to x^{(h)}(t)$是线性的,并关于给定的收敛连续.同样,物理学中所有的微分算子都是有关函数空间的连续线性算子,尽管或多或少是容易定义的.积分算子亦然.

一个基本的特例是把拓扑向量空间E映入实数集$\mathbf{R}$的连续线性映射$x(t)\to u(x)\in\mathbf{R}$,使得

$$u(x+y)=u(x)+u(y)$$

对x与y连续

$$u(\lambda x)=\lambda u(x)$$

对λ与x连续.

例子　设E是$[0,1]$上的连续数值函数$x(t)$的集合;在区间$[0,1]$上取定一数t_0.映射$x(t)\to x(t_0)$对每个函数给出它在t_0处的值,这个映射是连续线



性映射.

同样,对同一个 E,映射 $x(t)\to\int_0^1 x(t)\mathrm{d}t$ 是线性的,根据中值定理也是连续的.

更一般,设 $f(t)$ 是开始就取定了的数值连续函数,则映射 $x(t)\to\int_0^1 x(t)f(t)\mathrm{d}t$ 是连续线性映射.

同样的,如果与 $x(t)$ 相应的值是勒贝格 - 斯蒂尔吉斯积分,即 $x(t)\to\int_0^1 x(t)\mathrm{d}\mu(t)$,这里 $\mu(t)$ 是有界变差函数,则映射也是连续线性的.此外,同这些积分有关的那些定理及其逆定理使它们得到了现代的定义,即是定义为连续函数集合上的线性形式,对于一致收敛拓扑结构是连续的.

局部凸拓扑向量空间的汉恩 - 巴拿赫定理

这个定理断定了这种空间上存在连续线性形式,并明确提出确定这些形式的条件.

设 F 是 E 的向量子空间(即是对于加法运算以及用实数相乘的运算保持不变的子集合).因为 E 上存在拓扑结构,故可定义 F 在 E 中的闭包(或附贴包)$\bar{F}$.说明了这一点,就可以提出巴拿赫的下面这一条断言了:

给了闭向量子空间 $\bar{F}$ 和不在 $\bar{F}$ 中的一点 p,则 E 上存在一个连续线性形式,在集合 $\bar{F}$ 的每一点 x 处为零,而在 p 处不为零.

显然,若连续映射 $u(x)$ 在 F 上为零,则在 $\bar{F}$ 上也为零,因为 $\bar{F}$ 的点是 F 的点的极限,但不能由此推断 $\bar{F}$ 以外的任何情况.定理的目的就是说明 $u(x)$ 的存在.

第十五讲　维数的概念[①]

1　n 维流形

n“维”空间最著名的例子是数值空间(即“欧氏空间”)$\mathbf{R}^n$,它的点是 n 元实数组$(x_1,\cdots,x_n)$. $\mathbf{R}^n$ 的每一张超平面,即线性方程(不必齐次)$f(x)=0$所确定的集合 H,将空间 $\mathbf{R}^n$ 分成两个半空间,分别由 $f(x)>0$ 与 $f(x)<0$ 确定;又,任意给定两个不同的点 x 与 y,总存在超平面 H,将 x 与 y“分开”,即是 x 属于 H 所确定的一个半空间,而 y 属于另一个半空间,注意,H 同胚[②]于空间 $\mathbf{R}^{n-1}$,所以是 $n-1$“维”的.

从空间 $\mathbf{R}^n$ 出发,还可以造出另外一些空间,看来自然也应该叫做 n“维”空间. 例如,设 S_n 是空间 $\mathbf{R}^{n+1}$ 中由方程 $x_1^2+\cdots+x_{n+1}^2=1$ 确定的球面,这是一个有界闭子空间,所以是紧致空间;此外,球面投影在挖掉一点的 S_n 与空间 $\mathbf{R}^n$ 之间确立一同胚映射,所以 S_n 的每一点都有一个开邻域同胚于 $\mathbf{R}^n$.

再考虑(实)射影空间 P_n,它的点由 $n+1$ 个不全为零的齐次坐标 $Z_0,\cdots,Z_n$ 确定,我们规定:两组成比例的齐次坐标确定射影空间 P_n 的同一点;老实说,为了完全确定 P_n,必须确切说明 P_n 的拓扑结构如何,不过我们不详述了. P_n 中使得坐标 Z_i 为零的点组成的集合 H_i 是一个闭子集,同胚于 P_{n-1};子集 H_0 就是常说的“无穷远超平面”. H_i 的补集 U_i 是 P_n 的开集,同胚于 $\mathbf{R}^n$(例如,U_0 是 P_n 中所谓处于“有限距离”的点集):U_i 的点可以表为 n 个实数,即下面的商数

$$\frac{Z_0}{Z_i},\cdots,\frac{Z_{i-1}}{Z_i},\frac{Z_{i+1}}{Z_i},\cdots,\frac{Z_n}{Z_i}$$

因此,P_n 是 $n+1$ 个开集的并集,其中每一个开集都同胚于 $\mathbf{R}^n$.

一般,所谓 n 维流形,是指一个拓扑空间 V,满足豪斯道夫(Hausdorff)隔离公理(即:对每一对不同的点 x,y,存在 x 的一个邻域与 y 的一个邻域,没有公共点),而且具有下列性质:V 的每个点都有一个开邻域同胚于空间 $\mathbf{R}^n$. 如上所见,$\mathbf{R}^n$ 本身,S_n,P_n 都是 n 维流形.

---

① 原作者 H. Cartan(索尔本大学教授).

② 两个拓扑空间同胚是指其间存在一个同胚映射,即一个一一对应,使得开子集互相对应;也可以说,这个一一对应关系确定的把一个空间映成另一个空间的映射是双连续的.



问题 一个 n 维流形 V 与一个 m 维流形 W,当 $n \neq m$ 时,是否可能同胚?如果可能,那么流形的"维数"这个概念就没有任何拓扑含义.幸好,这是不可能的;不过其证明极难.

让我们更仔细地分析这个问题.整个问题就是说:当 $n \neq m$ 时,$\mathbf{R}^n$ 的开集 U 与 $\mathbf{R}^m$ 的开集 U' 之间是否存在同胚映射?(不言而喻,U 与 U' 假定非空)假设存在,则有把 U 映成 U' 的连续映射 f 以及把 U' 映成 U 的连续映射 g,使得复合映射 $g \circ f$ 是 U 的恒等映射,$f \circ g$ 是 U' 的恒等映射.若再假设 f 与 g 有连续偏导数,我们就可以证明这种同胚是不可能的:U' 中点的坐标 y_j 是 U 中点的坐标 x_i 的连续可微函数 f_j

$$y_j = f_j(x_1, \cdots, x_n)$$

同样

$$x_i = g_i(y_1, \cdots, y_m)$$

对上面两式微分得

$$\mathrm{d}y_j = \sum_i \frac{\partial f_j}{\partial x_i}\mathrm{d}x_i, \mathrm{d}x_i = \sum_j \frac{\partial g_i}{\partial y_j}\mathrm{d}y_j$$

以 $\frac{\partial f_j}{\partial x_i}$ 为元素的矩阵,在每一点 $x \in U$,定义了把 $\mathbf{R}^n$ 映入 $\mathbf{R}^m$ 的一个线性映射;以 $\frac{\partial g_i}{\partial y_j}$ 为元素的矩阵在对应点 $y = f(x)$ 处定义了把 $\mathbf{R}^m$ 映入 $\mathbf{R}^n$ 的一个线性映射.注意,$f \circ g$ 与 $g \circ f$ 是恒等映射,复合函数的导数定理表明:上述两个矩阵的积(次序任意)是单位矩阵.换言之,若 $x \in U$, $y = f(x)$,则有两个线性映射,一个把 $\mathbf{R}^n$ 映入 $\mathbf{R}^m$,另一个把 $\mathbf{R}^m$ 映入 $\mathbf{R}^n$,它们彼此互为逆映射;这表明,$\mathbf{R}^m$ 与 $\mathbf{R}^n$ 作为向量空间是同构的,但是代数学上的一条定理断言:$\mathbf{R}^n$ 中一组基底所含向量的个数 n 应等于 $\mathbf{R}^m$ 中一组基底所含向量的个数 m,此与假设 $n \neq m$ 不合.

我们刚才就连续可微的同胚映射证明了维数的不变性;为此,利用了向量空间"维数"的不变性,这里"维数"一词是就纯代数的意义而言(基底向量的个数).

要就任意同胚映射(不必可微)来证明 $\mathbf{R}^n$ 的维数不变性则困难得多.第一个证明(1911 年)属于布劳威尔(Brouwer).现在谈不上转引他的证明,也谈不上介绍证明的思路.他的证明基于单纯逼近的概念,与"同调"的概念密切相关.

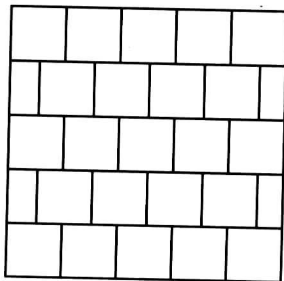


图 1

维数不变性的另一证明基于勒贝格(Lebesgue)的一个创造性思想①,也与同调有关.考虑 n 维方体 I^n,即线段 $I = [0,1]$ 自身的 n 次乘积.我们用一些闭集合把它铺满($n = 2$ 的情形见图 1),使得方体的每

① 有人说勒贝格在瓦兹(Oise)省古维欧这个小乡村的住宅里修了一堵砖墙,才想出了这样来说明维数的特性.

一点至多属于 $n+1$ 个这样的集合；利用直径任意小的集合，这总是可能的．此外，我们能证明，对于 I^n 的由充分小的闭集作成的任何覆盖，至少总有方体的一点，此点至少属于该覆盖的 $n+1$ 个集合．因此，数 n 说明空间 I^n 的一个纯拓扑性质，由此容易推出，只要 $m \neq n$，则不可能存在同胚映射，把 $\mathbf{R}^n$ 的开集映成 $\mathbf{R}^m$ 的开集．下面是勒贝格定理的确切陈述：若 I^n 由诸闭集 F_i 覆盖，使得每一 F_i 决不同时与方体的两个相对面相交，则 I^n 中有一点，同时属于 $n+1$ 个 F_i．

2　一种维数论的公理

我们但愿能谈谈比上面讨论过的更一般的空间的维数，例如，考虑 $\mathbf{R}^n$ 的一个子空间 A，为确定计设为闭子空间：我们能否对这样的空间定义一个维数，使得它是拓扑不变的？因此，把 A 换成 $\mathbf{R}^m$ 的子空间 B，只要 B 与 A 同胚，则 A 与 B 的维数一样．

在维数论中，我们应首先明确要定义维数的哪一类拓扑空间 C（维数是一个整数，可能为零或无限）．假定类 C 中只含有满足豪斯道夫隔离公理的空间，并包含所有的多面体．先说明一下所谓的多面体，为此要从 p 维单形讲起．在 $\mathbf{R}^n$ 中取 $p+1$ 个点 $x_i(i=0,\cdots,p)$，使得维数 $<p$ 的任何平面不可能包含所有这些点（当 $n \geqslant p$ 时，总是可能的），这些质点（质量可为0）的重心①的轨迹是包含这些点的最小凸集，我们将这样的 $p+1$ 个点的点组 $(x_0,x_1,\cdots,x_p)$ 称为以 $x_0,\cdots,x_p$ 为顶点的 p 维单形，它的每一点可唯一地表成 $\sum\limits_i \lambda_i x_i$ 诸实数 $\lambda_i \geqslant 0$，满足 $\sum\limits_i \lambda_i = 1$．具有同一维数 p 的两个单形同胚：$p=0$ 时是一个点，$p=1$ 时是一条线段，$p=2$ 时是一个三角形，$p=3$ 时是一个四面体②．于是，多面体是一个空间，可以被有限多个闭子空间所覆盖，这些闭子空间都与单形同胚（单形的维数可能不同），并且规则相处：覆盖中任两个单形的交是这两个单形的“面”（p 维单形的面是该单形的某些顶点的重心轨迹，这是一些维数为 $0,1,\cdots$，或 $p-1$ 的单形）．

---

① 假设在点 x_i 处有质量 $m_i \geqslant 0(i=0,1,\cdots,p)$，$\sum\limits_i m_i \neq 0$，则这 $p+1$ 个质点的重心是

$$x = \frac{\sum\limits_i m_i x_i}{\sum\limits_i m_i} = \sum_i \frac{m_i}{\sum\limits_j m_j} x_i, \lambda_i = \frac{m_i}{\sum\limits_j m_j}$$

称为这个质点系的重心坐标，合于 $\lambda_i \geqslant 0$，$\sum\limits_i \lambda_i = 1$．因此，这些点所能构成的一切质点系的重心轨迹是包含这 $p+1$ 个点的最小凸集．

② 这一段自此以下叙述不太清楚，改述如下：$\mathbf{R}^n$ 中的 p 维单形的同胚象称为 p 维拓扑单形，于是，多面体是一个拓扑空间，可以“剖分”为有限多个拓扑单形（维数可以各不相同），这些拓扑单形规则相处，即是任何两个拓扑单形的交要么是空集，要么是它们的公共“面”（p 维单形的面是指它的顶点子集构成的质点系的重心轨迹，所以是维数为 $0,1,\cdots$，或 $p-1$ 的单形；这些面的同胚象就是相应拓扑单形的面）．





现在回到类 C:我们要求它包含所有的多面体,此外还满足下述条件:与类 C 中某个空间 X 的闭子空间同胚的任何空间仍在 C 内.这样的类是相当广泛的,它包含了多面体的所有闭子空间.

类 C 的例子 1)紧致空间类 C_1(空间 X 是紧致的,如果它满足豪斯道夫公理,并且有波莱尔(Borel)-勒贝格性质:对于 X 的每个由开集族 $\{U_i\}$ 组成的覆盖,存在有限多个 U_i 覆盖 X).

2)局部紧致空间类 C_2(X 是局部紧致空间,如果它满足豪斯道夫公理,并且每一点至少有一个紧致邻域).空间 $\mathbf{R}^n$ 属于类 C_2,但不属于类 C_1.

3)可数型可尺化空间类 C_3,特别是对任何 n,这一类中含有 $\mathbf{R}^n$ 的所有子空间.空间 X 叫做可尺化的,如果 X 的拓扑结构可用距离定义,或者,更确切地说,用"尺度"定义.一个尺度对于任何两点 x,y,相应地给出一个数 $d(x,y)\geqslant 0$,使得

$$d(x,y)=0 \text{ 的充要条件是 } x=y$$
$$d(x,y)=d(y,x)$$
$$d(x,z)\leqslant d(x,y)+d(y,z) \text{(三角形不等式)}$$

这种可尺化的空间 X 是"可数型的",如果 X 中存在可数处处稠密集 $\{x_1,\cdots,x_n,\cdots\}$(即 X 的每一非空开集至少含其中一点).空间 $\mathbf{R}^n$ 属于 C_3:因为坐标为有理数的点构成可数处处稠密集,可以证明 C_3 类中的空间 X 的每一子空间 A 属于 C_3,即使 A 不是 X 的闭集.

4)正规空间类 C_4:空间 X 是正规的,如果它满足豪斯道夫公理,并且只要 A 与 B 是不相交的闭子集,就存在 X 上的数值连续函数,在 A 上为 0,在 B 上为 1.正规空间的每一闭子空间仍是正规的.

我们指出,每个紧致空间是正规的(换言之,C_4 包含 C_1);每个可尺化的空间是正规的(故 C_4 包含 C_3).

对于上述各类空间,都有一种维数理论.推而广之,对于维数论应该提出什么要求呢?我们假定,对于所考虑的类 C 的每一空间 X,都相应地有一整数 $\dim X$,称为 X 的维数,满足下列条件:

(Ⅰ)$\dim X\geqslant -1$;$\dim X$ 可以是无穷大;$\dim X\geqslant 0$ 的充要条件是 X 非空;

(Ⅱ)若 X 与 X' 同胚,则 $\dim X=\dim X'$;

(Ⅲ)若 A 是 X 的闭子空间,则 $\dim A\leqslant \dim X$;

(Ⅳ)若 X 是两个闭子空间 A_1 与 A_2 的并集,则

$$\dim X=\sup(\dim A_1,\dim A_2)$$

(Ⅳ′)(不必作为公理)若 X 是可数多个闭子空间 A_i 的并集,则

$$\dim X=\sup(\dim A_i)$$

(Ⅴ)方体 I^n 的维数是 n.

因此,若 $\mathbf{R}^n$ 在类 C 中,则 $\mathbf{R}^n$ 的维数是 n:事实上,由(Ⅲ),$\dim \mathbf{R}^n \geqslant \dim I^n$,故由(Ⅴ),$\dim \mathbf{R}^n \geqslant n$;另一方面,$\mathbf{R}^n$ 是可数多个闭方体的并集,故 $\dim \mathbf{R}^n \leqslant n$(根据 Ⅳ′).

注 必须避免提出下述公理:"若 f 是把 X 映入 Y 的连续映射,则象 $f(x)$ 的维数至多等于 $\dim X$."事实上,"皮亚诺(Peano)曲线"是把线段 I 映成正方形 I^2 的连续映射,所以这个条件与公理(Ⅴ)矛盾.

除上述公理外,还有一个性质迥异的公理.先引入一个概念:给了空间 X 的两个不相交的闭子集 A 与 B,我们说,闭子集 C 把 A 与 B 隔开,如果 $X-C$ 是两个分别包含 A 与 B 的不相交的开集的并集.新公理如下:

公理(A):若空间 X 中两个不相交的闭子集总可以被一个维数 $\leqslant n-1$ 的闭子空间隔开,则 $\dim X \leqslant n$.

3 门格-乌利松定理

我们取可数型可尺化空间类 C_3 作类 C,并且不仅要求公理(A),还要求其逆:若 $\dim X \leqslant n$,则 X 的两个不相交的闭子集总可以被一个维数 $\leqslant n-1$ 的闭子空间隔开.这个加强了的公理与公理(Ⅰ)完全确定了维数.确切地讲,用关于 n 的归纳法,定义什么是维数 $\leqslant n$ 的空间 X,按定义,这就是指 X 的任意两个不相交的闭子集都可以被一个维数 $\leqslant n-1$ 的闭子空间隔开.因为根据公理(Ⅰ),已经知道了维数 $\leqslant -1$ 的空间(即唯一存在的空集空间),从而依次得到维数 $\leqslant 0$ 的空间,维数 $\leqslant 1$ 的空间等等的定义.于是空间 X 的维数等于满足 $\dim X \leqslant n$ 的最小整数 n;如果不存在这样的整数 n,就说空间 X 的维数是无穷大.

这样定义的维数,不曾考虑到条件(Ⅱ)至(Ⅴ).事实上,它们都起作用(包括(Ⅳ′)在内),现在是一些定理了;为证明它们(往往很难),要用到空间 X 是可数型可尺化空间这个事实.还有,公理(Ⅲ)对于 X 的任何子空间 A,即使是非闭子空间都成立.

有关这种维数论的问题,读者可参考胡尔维兹(Hurewicz)-瓦尔曼(Wallman)的著作(《Dimension Theory》,Princeton, 1941),或者法瓦德(Favard)的小册子(《Espaceet Dimension》,Albin Michel, 1950).

对于 C_3 类,如果有另一套维数论,则空间 X 按照这套新理论的维数至多等于它按照门格-乌利松理论的维数:证明用公理(A)与归纳法.

应该阐明零维空间是什么.我们指出,X 是 0 维空间的充要条件是:X 的每





个点都具有基本邻域组,使得邻域的边界①是空集.然而,边界是空集的集合不过就是既开又闭的集合,0 维空间的例子:设 D 是区间 I 的无理点所成的子空间,D 是 0 维空间,因为 D 的每一点 x 都有一族既开又闭的基本邻域,即区间 I 中含有 x 且具有有理端点的那些子区间与 D 的交.存在由无限多点构成的 0 维紧致空间:例如,由数字 0,1 构成的无限排列的空间 E,其拓扑结构的定义是:$S_1,\cdots,S_k,\cdots$ 这些排列以排列 S 为极限的充要条件是:对于每一整数 n 存在一整数 $k(n)$,使得 $k > k(n)$ 时,排列 S_k 与 S 中头 n 个数字都相同.这个空间 E 是紧致的,维数是0(头 n 个数字相同的所有排列构成 E 中既开又闭的集合).存在一个连续映射 f,把这个空间 E 映成区间 $I = [0,1]$;对于0与1组成的任何排列 $(a_1,\cdots,a_k,\cdots)$,f 相应地给出实数 $\sum_k \frac{a_k}{2^k}$(此级数收敛),这个数以所给排列为其二进位表示,除非从某位开始数字都是 1;在这种情形下,例如 0101111… 这个排列所确定的数等于排列 0110000… 所确定的数.因此把 E 映成 I 的这个连续映射 f 不是一一的.

4　亚力山大洛夫 – 捷赫理论

这种理论是对勒贝格的思想(见上面的 2) 加以适当修改建立起来的.我们取正规空间类 C_4 为类 C.给了空间 X 的一个覆盖,由有限个开集 U_i 组成,我们说这覆盖有维数 $\leqslant n$,如果 X 的每一点至多属于覆盖的 $n+1$ 个集合.另一方面,我们说,开集 V_j 组成的覆盖比开集 U_i 组成的覆盖更细,如果每个 V_j 至少含于一个 U_i 中.在亚力山大洛夫 – 捷赫的理论中,有如下定义:

正规空间 X 有维数 $\leqslant n$,如果对 X 的每一个由有限多个开集组成的覆盖,存在一个更细的覆盖,也由有限多个开集组成,并且维数 $\leqslant n$.

我们指出,这样定义的维数满足公理(Ⅰ)到(Ⅴ)(可能除去(Ⅳ′)),以及公理(A).

已经说过,类 C_4 包括类 C_3.值得注意的是,在类 C_3 上,亚历山大洛夫 – 捷赫意义上的维数等于门格 – 乌利松意义上的维数.

另一方面,类 C_4 包括紧致空间类 C_1,因此亚历山大洛夫 – 捷赫的理论也就产生了紧致空间的维数论(门格 – 乌利松的理论只给出了可尺度化的紧空间的维数).紧致空间的这种维数理论满足前述所有公理,包括(Ⅳ′).

现在我们要给出紧致空间维数的另一特征.有一类紧致空间,即多面体,其

① 对于拓扑空间 X 的一个子集 A,可以相应地得到下面两个集:付贴包(或闭包)$\bar{A}$,这是包含 A 的最小闭集;内部 A,这是含于 A 中的最大开集.差集 $\bar{A} - \mathring{A}$ 称为 A 的边界,这是一个闭集,也是 A 的补集的边界.A 的边界是空集,就是指 $\mathring{A} = \bar{A}$,也就是说 A 既是开集又是闭集.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

维数是明显的.事实上,根据公理(Ⅰ)到(Ⅴ),一个多面体的维数显然是拼成该多面体的诸单形的最大维数.这表明紧致空间维数的概念可归结为多面体维数的概念.为简便计,假设所考虑的紧致空间 X 的拓扑结构由距离确定;依定义,把 X 映入一多面体 P 的连续映射是 ε 映射,如果 P 的每一点的逆象是 X 中直径 $\leqslant \varepsilon$ 的集合(这里 ε 表示一个正数).我们指出下列定理:紧致空间 X 有维数 $\leqslant n$,必须且只需对于每一 $\varepsilon > 0$,存在维数 $\leqslant n$ 的多面体 P 以及把 X 映入 P 的 ε 映射.

下面是这个准则的一个应用.几乎显然:两个多面体 P 与 Q 的积能剖分成单形,如此剖分后的 $P \times Q$ 的维数是 P 与 Q 维数的和,由此推出,若 X 与 Y 是两个紧致空间,则有

$$\dim (X \times Y) \leqslant \dim X + \dim Y \tag{1}$$

遗憾的是,等式不总是成立:邦特列雅金在1930年给出一例,两个维数为2的紧致空间,其积的维数是3.

不等式(1)在亚历山大洛夫－捷赫的理论中也真.事实上,若 X 与 Y 属于类 C_4,则可将它们扩张成紧致空间 $\tilde{X}$ 与 $\tilde{Y}$,使得 $\dim \tilde{X} = \dim X, \dim \tilde{Y} = \dim Y$,于是 $X \times Y$ 扩张成 $\tilde{X} \times \tilde{Y}$,并有

$$\dim (X \times Y) \leqslant \dim (\tilde{X} \times \tilde{Y}) \leqslant \dim \tilde{X} + \dim \tilde{Y} = \dim X + \dim Y$$

这就证明了(1).

5 局部紧致空间的维数

在这一理论中,我们将局部紧致空间 X 的维数定义成 X 的紧致子空间维数的上确界.于是得到了类 C_2 的维数理论,它满足(Ⅰ)至(Ⅴ)的所有公理(包括(Ⅳ′)),也满足公理(A)与关系(1).

因此,我们实际上有两种维数理论:一种是空间类 C_3 的维数论,现在这种是局部紧致空间类 C_2 的维数论.使人愉快的是:在公共类 $C_3 \cap C_2$ 上,两种理论一致:若一局部紧致空间 X 属于类 C_3(为此,必须且只需 X 是可数多个可尺度化的紧致子空间的并集),则 X 在门格－乌利松意义上的维数正好等于 X 中紧致子空间维数的上确界,这是因为门格－乌利松的维数满足公理(Ⅳ′).

上述两种理论中,维数的概念都有一种局部特征:$\dim X \leqslant n$ 的充要条件是:每一点都有一个闭邻域,其维数 $\leqslant n$.

6 空间 $\mathbf{R}^n$ 的子空间



设 X 是 $\mathbf{R}^n$ 的子空间,故 X 属于类 C_3.若 X 包含 $\mathbf{R}^n$ 的一个非空开集,则显



然 $\dim X = n$.其逆亦真:若 X 是 $\mathbf{R}^n$ 的子空间,维数恰为 n,则 X 包含 $\mathbf{R}^n$ 的一个非空开集.在 X 是紧致子空间的特别情形下,这一点是容易理解的:事实上,假设 X 没有内点,取多面体 $P \subset \mathbf{R}^n$,它包含 X 并且由直径 $\leqslant \varepsilon$ 的单形组成.在 P 的每个 n 维单形内部选取一点不属于 X,从这点把 X 在这个单形中的那部分投影到单形的边界上,对于 P 的所有 n 维单形如法炮制,便得一个连续映射,把 X 映入 P 中维数 $\leqslant n-1$ 的那些单形的并集 P',这是一个 ε 映射,把 X 映入一个维数 $\leqslant n-1$ 的多面体,所以 $\dim X \leqslant n-1$.

紧致空间是些什么样的空间呢?紧致空间可以实现为欧氏空间 $\mathbf{R}^q$ 的闭子空间(因而也可以实现为方体 I^q 的子空间).这样的空间显然应是可尺度化的.其逆定理亦真(门格 - 诺贝宁):若 X 是 n 维可尺化的紧致空间,那么 X 可以实现为 $2n+1$ 维方体 I^{2n+1} 的闭子空间.但是,的确有一些 n 维可尺度化的紧致空间,不能实现为方体 I^{2n} 的子空间(例如,$2n+2$ 维单形的诸 n 维面的并集).

7 映入球面 S_n 的映射

设 X 是紧致空间,A 是 X 的闭子空间;把 A 映入空间 $\mathbf{R}^n$ 的每一连续映射可扩张成把整个 X 映入 $\mathbf{R}^n$ 的连续映射(当 X 为正规时也真).但是,如果考虑映入球面 S_n 的映射,则结论不真:例如,设 B_{n+1} 是 $n+1$ 维球体,即 $\mathbf{R}^{n+1}$ 中到原点的距离 $\leqslant 1$ 的点集(S_n 是 B_{n+1} 的边界),f 是 S_n 的恒等映射,则可证明,f 不能扩张成把 B_{n+1} 映入 S_n 的连续映射.

紧致空间的维数有如下特性:$\dim X \leqslant n$ 的充要条件是:对任意闭子空间 $A \subset X$,把 A 映入 S_n 的连续映射 f,可以扩张成把 X 映入 S_n 的连续映射.

对于门格 - 乌利松的维数(如果 X 是可数型可尺度化空间),上述性质也成立.

由此推出下述性质:若 Y 是局部紧致空间 X 的闭子空间,则有

$$\dim X \leqslant \sup(\dim Y, \dim(X-Y)) \tag{2}$$

事实上,在门格 - 乌利松理论中也有上面的不等式,因为这时开集 $X-Y$ 是可数多个闭集的并集.

8 补子空间

空间 X 叫做连通的,如果它不是两个不相交的非空开子集的并集.如果从连通空间 X 中除去一闭子集 A,则补空间 $X-A$ 可能不连通(例如:X 是空间 $\mathbf{R}^n$,A 是超平面),然而,我们指出,若 X 是 n 维连通流形,A 是维数 $\leqslant n-2$ 的闭子空间,则空间 $X-A$ 连通(例如,除去一点的平面是连通的),当 A 为 $n-1$ 维

时,确切地讲,A 作为拓扑空间是 $n-1$ 维流形时,结果又如何呢?我们不加证明地提出一般的约当 - 布劳威尔定理:设 X 是 n 维"可定向"流形,连通而且"单连通"(这就是说,每条闭曲线可形变为一点),若一个 $n-1$ 维流形(不能是否连通)嵌入 X 成为闭子空间 A,则 $X-A$ 的连通区的个数等于 A 的连通区个数加 1;此外,A 的每一连通区必是"可定向的".这个定理特别适用于 X 是空间 $\mathbf{R}^n$ 的情形.

9 与 k 维测度的关系

设 X 是可数型可尺化空间,在 X 上确定了一尺度,于是,对每个整数 n,令 $m_n(X)=\inf\limits_{\{X_i\}}\sum\limits_i d(X_i)^n$,这里诸 X_i 组成 X 的可数覆盖,对所有这种覆盖求下确界,$d(X_i)$ 表示按已知尺度算出的 X_i 的"直径";当 X 紧致时,可以只考虑有限覆盖.我们指出:若 $m_{n+1}(X)=0$,则 $\dim X\leqslant n$.有一个类似于逆定理的结果:若 $\dim X\leqslant n$,则 X 上存在一个尺度,使得 $m_{n+1}(X)=0$;还可以把 X 实现为 I^{2n+1} 的子空间,使得方体的尺度产生的 X 的尺度就合乎要求.

例如,取区间 I 作为 X,这是 1 维空间,区间内两点间的通常距离作为尺度.显然,I 可以被有限多个区间覆盖,这些区间长度的平方和可以任意小.

结论 我们希望,上面这个硬性压缩的概述能够使读者对拓扑学中与"维数"概念有关的种种问题有所了解.

附录:皮亚诺曲线

考虑等腰直角三角形 T,设 S_0 与 S_1 是两锐角的顶点.我们来定义一个连续映射 g,把区间 $I=[0,1]$ 映成 T,使得 $g(0)=S_0$,$g(1)=S_1$.将这样得到的"曲线"再添上关于 T 的斜边的对称图形,我们就得到一条填满正方形的"曲线".

为此,再考虑前面提到的空间 E.现在定义把 E 映成 T 的连续映射 h.T 的直角平分线将 T 分成两个(闭)三角形 T_0 与 T_1,T_0 含有顶点 S_0,T_1 含有顶点 S_1.同样,T_0 是两个相等三角形 T_{00} 与 T_{01} 的并,T_1 是两个三角形 T_{10} 与 T_{11} 的并,这里 T_{00} 是含有 S_0 的三角形,T_{01} 是含有 T 的直角顶点的三角形.同样,T_{10} 含有 T 的直角顶点,T_{11} 含有 S_1.重复这个过程,把三角形 T_{00},T_{01},T_{10},T_{11} 的每一个又各分成两个相等的三角形,用 T_{000},T_{001} 表示 T_{00} 所分成的两个三角形,等.确切地讲,诸三角形的编号使得序列

$$T_{000},T_{001},T_{010},T_{011},T_{100},T_{101},T_{110},T_{111}$$

中,相继的两个三角形恒有一条公共边(图 2).如此继续下去;对于每个整数 n,



三角形 T 被分成 2^n 个相等的三角形,其中每一个对应于由 0 或 1 组成的 n 个数的排列.

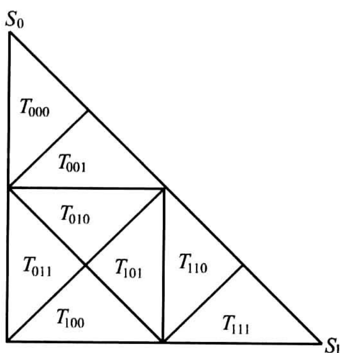


图 2

现在考虑空间 E 的一点,即数字0或1的一个无限排列,例如,0010111…;对于这个排列,我们相应地得到三角形的无限序列

$$T_0, T_{00}, T_{001}, T_{0010}, T_{00101}, T_{001011}, T_{0010111}, \cdots$$

其中每一个都包含下一个,它们的直径趋于零,所以有唯一的公共点.若用 $h(x)$ 表示这一公共点,就定义了一个把 E 映入 T 的映射 h,可验明 h 连续.因为 T 的每一点至少属于三角形套的无限序列,故 h 把E 映成T.然而,前面3曾定义了把 E 映成区间 $I=[0,1]$ 的连续映射 f;另一方面,若两个无限排列(例如 $x=0101111\cdots$ 与 $y=0110000\cdots$)使得 $f(x)=f(y)$,则易证 $h(x)=h(y)$,即确定 T 的同一点.若对每一点 $u\in I$,选取一 $x\in E$,使得 $f(x)=u$,则元素 $h(x)\in T$ 与x 的选取无关,记为 $g(u)$,由此得到把 I 映成T 的映射g,g 是连续的,因为它把区间$\left[0,\frac{1}{2}\right]$映成 T_0,把区间$\left[\frac{1}{2},1\right]$映成 T,把$\left[0,\frac{1}{4}\right]$映成 T_{00},等.这映射 g 正是要求的"皮亚诺曲线".

注意,对于任何一点 $t\in T$,至多存在 I 的四个点,被 g 映成t;一般只有一个这样的点.

图 3 是皮亚诺曲线的近似(画出的三角形有六个数字的脚标).

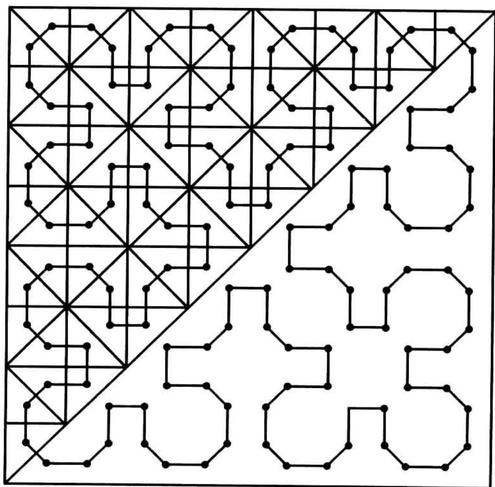


图 3

第十六讲　覆叠与基本群①

1　引　言

覆叠的概念可以按照几何问题中经常碰到的一种情况提出来;我们先讲几个例子:

a) 设 Y 是圆柱体上的螺旋线(图 1),用 p 表示把螺旋线 Y 映成圆柱体底圆 X 的投射算子:$Y\to X$.这个映射是局部同胚映射:若 y 是 Y 的一点,$x=p(y)$ 是其投影,则在 y 的适当邻域(例如一小段弧)与 x 的适当邻域之间,p 定义了一个一一且双连续对应.不言而喻,p 不是整体同胚映射,因为 Y 的相异点在 X 内可以有相同的象.我们说 Y 是 X 的无限叶覆叠.

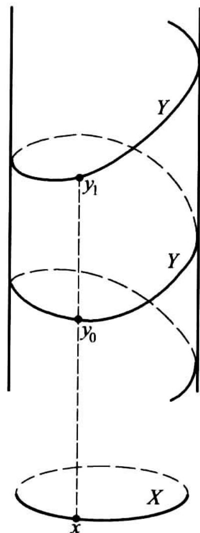


图 1

换一种提法,可以把上述覆叠叫做"角度"覆叠.对于每个元素 $\alpha\in\mathbf{R}$,相应地得到角 α 在圆周 S_1 上确定的点 $p(\alpha)$.

b) 取空间 $\mathbf{R}^{n+1}$ 的球面 S_n 为 Y,即满足 $\sum x_i^2=1$ 的点 $(x_0,\cdots,x_n)$ 的集合;这时空间 X 是 n 维实射影空间,记为 P_n,投射 $p:Y\to X$ 对于点 $(x_0,\cdots,x_n)\in S_n$,相应地给出 P_n 中齐次坐标为 $(x_0,\cdots,x_n)$ 的点.这投射是把 Y 映成 X 的连续映射,X 的每一点都是 Y 的两个对径点的象.这是两叶覆叠的例子.

P_n 以 S_n 为覆叠这一事实可以用来研究 P_n 上的椭圆型非欧几何,相当于 S_n 中通常的几何结构(由 $\mathbf{R}^{n+1}$ 的尺度导出的几何结构).

c) 取 $Y=S_3$,视为范数是 1 的四元数 $a+bi+cj+dk$ 的集合,四元素的乘法使 S_3 具有拓扑群的结构;此外,如 E · 嘉当所证,这是除 S_1 外唯一具有这种性质的球面.

我们把三维空间的点 $M=(x,y,z)$ 与"纯"四元素 $xi+yj+zk$ 等同.若

① 原作者 J. P. Serre (法兰西学院教授).



$q \in S_3$，令 $R_q(M) = qMq^{-1}$，这里的乘积是四元数代数中的积；立即可证，$R_q(M)$ 仍是纯四元数，故为 $\mathbf{R}^3$ 的一点．还有，$R_q(M)$ 的范数等于 M 的范数．因此，把 $\mathbf{R}^3$ 映入自身的线性变换 R_q 保持距离不变；若用 $SO(3)$ 表示三维空间的旋转群，那么 $q \to R_q$ 是一连续同态映射 $p: S_3 \to SO(3)$．容易验明 p 是局部同胚，每个旋转 $\mathbf{R} \in SO(3)$ 都由两个互逆的四元数产生．本例又是两叶覆叠，可以使对旋转群的研究化为对 S_3 的研究，一般说来，这样更为简单．

比较例子 b) 与 c) 可知：群 $SO(3)$ 与射影空间 P_3 同胚，这个结果可以直接由几何论证得到．

2　覆叠的定义

设 X 与 Y 是两个拓扑空间，$p: Y \to X$ 是把 Y 映成 X 的连续映射．我们说，Y 是 X 由投影 p 确定的覆盖，如果对于每个 $x \in X$，存在 X 的开集 U，含有 x，具有下列性质：

(R) U 在 Y 中的逆象 $p^{-1}(U)$ 是一些互不相交的开集 V_i 的并集，使得对于每个指数 i，投影 $p: V_i \to U$ 是把 V_i 映成 U 的同胚映射(图 2)．

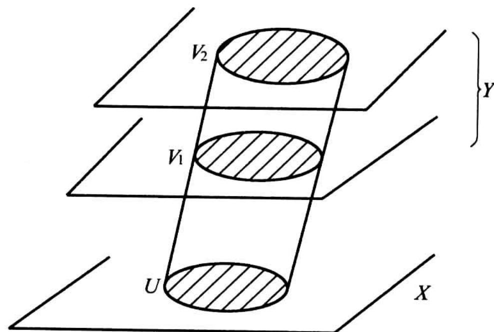


图 2

这个条件可以表达得更简单，只要我们先定义空间 X 的所谓简单覆叠：这是指积空间 $X \times E$，这里 E 是一个离散空间，映射 $p: X \times E \to X$ 不过是在第一个因子上的投影．于是条件(R) 可表达成：$P^{-1}(U)$ 同构于 U 的简单覆叠．换言之，每个覆叠局部同构于简单覆叠．(注意，如果不假设 E 离散，那就得更一般的纤维空间的概念)

不难验明，引言中诸例满足条件(R)．例 a) 中，只需取 U 为含有已知点 x 的任意圆弧．例 b) 中，取 U 为 P_n 中挖掉一张不通过 x 的超平面后的点集．

3 覆叠上的道路

覆叠的理论只有对空间 X 与 Y 作了"局部正则"的假设后才是令人满意的；假如把全不连通空间，例如康托完全集，取作空间 X，那会是荒唐可笑的．我们不打算提出最少的假设；今设（事实上这是过分限制了）X 是拓扑流形．这就是说 X 是隔离空间，它的每一点都有与 $\mathbf{R}^n$ 的开集同胚的开邻域．再设 X 是连通的；这时，整数 n 是常数，即空间 X 的维数（参见 H·嘉当的讲演）．由于条件（R），Y 也是 n 维流形（但不必连通）．

直观上看，空间 X 与 Y 显然不能太不相同．它们是局部同构的（因此覆叠的概念也同微分几何有关），所以只有整体性质上的差异．我们可以设想，在空间 X 与 Y 中给出一个，看看有什么办法可以确定另一个．事实上，只要从 X 出发，便可找到简单的结论，这有赖于考虑 X 与 Y 上的道路．

按定义，拓扑空间 X 上的一条道路是一个连续映射 $f: I \to X$，这里 I 表示区间 $[0,1]$．如果用 x_t 代替 $f(t)$，那么一条道路就是 X 的一族点 x_t，连续依赖于下标 $t \in [0,1]$．点 x_0 与 x_1 分别称为道路的起点与终点；一条闭合的道路（即 $x_0 = x_1$）称为（在点 x_0 的）闭路．因 X 连通，所以给了起点和终点，总有一条连接它们的道路．

Y 上的一条道路 y_t 称为 X 上的一条道路 x_t 的提升，如果对每个 $t \in I$ 有 $p(y_t) = x_t$．这样的道路恒存在，确切地讲有：

若 x_t 是 X 上的一条道路，y_0 是 Y 的一点，被投射成 x_0，则 Y 上存在唯一一条道路，以 y_0 为起点并且是 x_t 的提升．

事实上，先设覆盖结构 $p: Y \to X$ 是简单的，即 $Y = X \times E$，E 离散，于是给定的 y_0 可以写成 $y_0 = (x_0, z)$，$z \in E$，所求的道路必为 $y_t = (x_t, z)$，特殊情形得证．一般情形可归结为刚才的特例，只要注意我们可把区间 $[0,1]$ 分割成一些区间 $[t_i, t_{i+1}]$，使得 $t \in [t_i, t_{i+1}]$ 时，x_t 含于一个开集 U_i，而 U_i 有性质（R），把刚才得到的结果应用到简单覆叠结构 $p^{-1}(U_i) \to U_i$，知：只要给了起点 y_{t_i}，则在 $t_i \leqslant t \leqslant t_{i+1}$ 时，存在唯一的 y_t．由此一段一段地得到 y_t 在 $0 \leqslant t \leqslant 1$ 时的存在及唯一性．

注意，若 x_t 是闭路，则道路 y_t 的终点 y_1 被投射成 x_0；但一般没有 $y_0 = y_1$，换言之，一条闭路的提升未必是闭路（图 3）．在例 a）中，若取圆周本身（具有适当的参数表示）为闭路 x_t，则点 y_1 位于 y_0 的"上面一圈螺旋线"上．在例 c）中，若取绕 z 轴转 $2\pi t$ 的旋转为 x_t，则有 $y_t = \cos \pi t + k\sin \pi t$，由此 $y_0 = 1$，$y_1 = -1$．这差不多就是覆叠结构的特性；为表明此点，需要基本群的概念，下面就来定义．



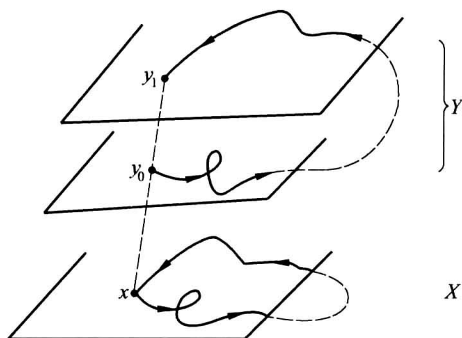


图 3

4　同伦的道路·基本群

设 X 是拓扑空间，f 与 g 是 X 上的两条道路，有相同的起点 a 与终点 b. 我们说，f 与 g 同伦，如果存在 X 的一族道路 f_u, $0 \leqslant u \leqslant 1$，都以 a 为起点 b 为终点，使得 $f_0 = f$, $f_1 = g$，并且连续依赖于 u（即 $f_u(t)$ 是偶对 (u,t) 的连续函数）. 易见就具有已知起点与终点的道路而言，同伦是一等价关系，记成 $f \sim g$.

同伦关系当然适用于在定点 x_0 的闭路. 在闭路之间还有一个合成法则：若 f 与 g 是两条闭路，它们的积 $f * g$ 是一条闭路，先走遍 f，然后再走遍 g 而得. 当然应当精确说明所用的参数表示

$$f * g(t) = \begin{cases} f(2t), & \text{若 } 0 \leqslant t \leqslant \frac{1}{2} \\ g(2t-1), & \text{若 } \frac{1}{2} \leqslant t \leqslant 1 \end{cases}$$

容易验明，这合成法则与同伦定义的等价关系是相容的，换言之有

$$f \sim f', g \sim g' \Rightarrow f * g \sim f' * g'$$

因此，在 x_0 的闭路（同伦）类的集合 $\pi_1(X, x_0)$ 中可以引进合成法则，即是由 $f * g$ 求商而得，从而 $\pi_1(X, x_0)$ 构成一个群，因为：

1）存在单位元素 e，即常值闭路的同伦类：对任何 t 有 $e(t) = x_0$.

应当验明：对每条闭路 f，我们有 $f \sim f * e$. 为此定义一族闭路 g_u, $0 \leqslant u \leqslant 1$，如下

$$g_u(t) = \begin{cases} f((1+u)t), & \text{若 } 0 \leqslant t \leqslant \frac{1}{1+u} \\ x_0, & \text{若 } \frac{1}{1+u} \leqslant t \leqslant 1 \end{cases}$$

我们有 $g_0 = f$, $g_1 = f * e$，由此 $f \sim f * e$. 同样可证 $f \sim e * f$.

2) 对每条闭路 f, f 的同伦类有逆元素,即是 $f'(t) = f(1 - t)$ 确定的闭路 f' 的同伦类.

应当验明: $f * f' \sim e$. 为此,定义一族闭路 h_u, $0 \leqslant u \leqslant 1$, 如下

$$h_u(t) = \begin{cases} f(2ut), & 若 0 \leqslant t \leqslant \frac{1}{2} \\ f(2u(1 - t)), & 若 \frac{1}{2} \leqslant t \leqslant 1 \end{cases}$$

我们有 $h_0 = e$, $h_1 = f * f'$, 由此 $f * f' \sim e$.

3) $\pi_1(X, x_0)$ 的合成法则是结合的.

应当验明:若 f, g, h 是三条闭路,则有

$$(f * g) * h \sim f * (g * h)$$

为此,定义一族闭路 k_u, $0 \leqslant u \leqslant 1$, 如下

$$k_u(t) = \begin{cases} f\left(\dfrac{4t}{1 + u}\right), & 若 0 \leqslant t \leqslant \dfrac{1 + u}{4} \\ g(4t - 1 - u), & 若 \dfrac{1 + u}{4} \leqslant t \leqslant \dfrac{2 + u}{4} \\ h\left(1 + \dfrac{4(t - 1)}{2 - u}\right), & 若 \dfrac{2 + u}{4} \leqslant t \leqslant 1 \end{cases}$$

我们有 $k_0 = (f * g) * h$, $k_1 = f * (g * h)$. 由此得到所要的结果.

当 X 是弧连通空间时(连通流形就是如此),可以证明,诸群 $\pi_1(X, x)$, $x \in X$, 都同构. 于是可以写 $\pi_1(X)$ 代替 $\pi_1(X, x_0)$, 称为空间 X 的基本群(或者普昂卡雷群). 当 $\pi_1(x) = 0$, 即每一闭路同伦于常值闭路时, X 叫做单连通空间.

除 $\pi_1(X)$ 外,可定义群 $\pi_n(X)$, $n \geqslant 2$, 称为 X 的同伦群,与覆叠理论无关.

5 底空间的覆叠空间的分类

继续前面的讨论,设 Y 是拓扑流形 X 的覆叠空间. 已经知道, X 上有同一终点的两条道路 x_t 与 x'_t 未必能提升成 Y 上有同样性质的道路;然而 x_t 与 x'_t 同伦时,确是如此. 当这两条道路足够接近时,结论显然,而一般的情形,可以化为上述情形.

于是,设 x_0 是 X 中的定点,令 $E = p^{-1}(x_0)$ 是 Y 中被投射成 x_0 的点 y 的集合. 若 $y \in E$, $f(t) = x_t$ 是 X 上在 x_0 处的闭路,则据前述, Y 上有唯一一条道路以 y 为起点,并且是道路 f 的提升,它的终点 y_1 只依赖于闭路 f 的同伦类 α, 因此,可表成 $y * \alpha$. 我们有下列等式

$$\begin{cases} y * e = y \\ y * (\alpha * \beta) = (y * \alpha) * \beta, 若 \alpha, \beta \in \pi_1(X) \end{cases}$$



对于每个$\alpha \in \pi_1(X)$，映射：$y \to y * \alpha$是集合E的置换，它对于两个元素α与β的积，相应地给出置换的积(顺序相反)．这时就说，我们得到$\pi_1(X)$由E的置换给出的表示．

底空间X的覆叠分类定理可以叙述如下：

$\pi_1(X)$由集合E的置换给出的每一表现对应于唯一一个覆叠$Y \to X$(当然是在同构的意义下唯一)．

只介绍证明的原理：总的说来，问题在于由X，E以及运算$y * \alpha$，$y \in E$，$\alpha \in \pi_1(X)$重新造出Y．令$z \in Y$，我们可用一条道路f把z在X中的投影$p(z)$与点x_0联结起来，求这条道路的提升，则至少有一条道路可以把z联结到一点$y \in E$．再有，知道了y与f便确定了z；两对元素(y,f)与(u,g)对应于同一点z的充要条件是：闭路$f * g'$的同伦类α把y变成u，即$u = y * \alpha$(图4)．因此，如果在元素对(y,f)的集合中引进下列等价关系

$$(y,f) \sim (u,g)，若\ u = y * (f * g')$$①

则相应的商集可以视为Y．

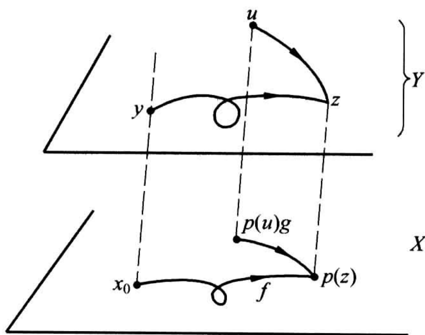


图4

剩下的事就是只有X，E以及运算$y * \alpha$来定义Y的拓扑结构．这不会出现严重的困难(不过正是这里需要X的局部正则性假设)．

最有趣的是Y为连通空间的情形，就是说E的两点可以用一条道路联结，或者说由$\pi_1(X)$确定的置换群是可迁的．若y_0是E内一定点，$\pi_1(X)$中使得$y_0 * \alpha = y_0$的元素$\alpha \in \pi_1(X)$的集合是$\pi_1(X)$的子群H，E可以视为$\pi_1(X)$的模H等价类的集合．因此我们看到X的连通覆叠与$\pi_1(X)$的子群成对应．特别，要使X的连通覆叠只是它自身(一叶覆叠)，充要条件是：X是单连通的．

若取H是$\pi_1(X)$的单位元所成的子群，那么得到的X的覆叠$\tilde{X}$称为X的万用覆叠．术语"万用"意味着$\tilde{X}$是X的任何连通覆叠Y的覆叠；换言之，投射p：

① 这里的$(f * g')$应该是它的同伦类．—— 译者注

$\tilde{X} \to X$ 被分解成 $\tilde{X} \to Y \to X$.

特别, $\tilde{X}$ 是单连通空间,这是 X 的唯一一个具有上述性质的连通覆盖.

若 Y 是 X 的一个覆叠,投射为 p, Y 的自同构是指任何连续映射 $g: Y \to Y$,满足 $p \circ g = p$,这样的映射必为把 Y 映成自身的同胚映射. 在 $Y = \tilde{X}$ 的特殊情形,我们指出 $\tilde{X}$ 的自同构群与 $\pi_1(X)$ 同构(同构映射与点 $y_0 \in \tilde{X}$ 的选取有关); $\tilde{X}$ 的两点 y_1 与 y_2 有同样的投影,充要条件是:存在 $\tilde{X}$ 的自同构 g,使得 $g(y_1) = g(y_2)$.

6 基本群的确定

我们已经看到,知道了空间 X 的基本群 $\pi_1(X)$,就可以确定 X 的所有覆叠. 因此,对于给定的空间 X,算出 $\pi_1(X)$ 是有益的. 下面举几个计算 $\pi_1(X)$ 的例子:

1) $X = \mathbf{R}^n$, n 维欧氏空间.

设 f 是起点与终点均为 $(0,\cdots,0)$ 的闭路. 若 $0 \leqslant u \leqslant 1$,令 $f_u(t) = uf(t)$,则 $f_1 = f$,而闭路 f_0 是常值闭路. 因此,任何闭路都同伦于0,这表明 $\mathbf{R}^n$ 单连通,换言之 $\pi_1(\mathbf{R}^n) = 0$.

更一般地,同样的推理也适用于 $\mathbf{R}^n$ 的每一凸开集.

2) $X = S_1$,圆周.

因 $\pi_1(\mathbf{R}) = 0$,故例 a) 的覆盖 $\mathbf{R} \to S_1$ 是 S_1 的万用覆叠,它的自同构群与 $\pi_1(S_1)$ 同构. 这自同构群显然由幅度为 na 的垂直移动构成,这里 a 是两圈螺线的距离,所以这个群同构于整数群 z,由此得到

$$\pi_1(S_1) = Z$$

3) $X = S_n$, $\mathbf{R}^{n+1}$ 中的 n 维球面(设 $n \geqslant 2$).

首先指出, S_n 上的每条闭路 f 可形变成一条"折线"型闭路,由首尾相接的大圆弧组成(把区间 $I = [0,1]$ 分成子区间,使得每个子区间上点 $f(t)$ 保持在一个固定半球的内部,再定义一个从这个半球到大圆弧上的形变). 然后,我们选取一点 P,不在新闭路 f' 上(对于 f 这不一定可能:皮亚诺曲线). P 在 S_n 中的补集与 $\mathbf{R}^n$ 同胚,作球面投影便可证实此点;根据1),闭路 f' 在这补集中与0同伦,自然在 S_n 中也同伦于0,所以 f 也同伦于0. 由此得到

$$\pi_1(S_n) = 0 \quad n \geqslant 2$$

4) $X = P_n$, n 维射影空间.

对于 $n = 1$, P_1 与 S_1 同胚,故 $\pi_1(P_1) = Z$. 现设 $n \geqslant 2$,这时例 b) 中的覆叠 $S_n \to P_n$ 是 P_n 的万用覆叠,因为 $\pi_1(S_n) = 0$;于是这个覆叠的自同构群与



$\pi_1(P_n)$ 同构.又因是两叶覆叠,故 $\pi_1(P_n)$ 是两个元素的群.由此得到

$$\pi_1(P_n) = \frac{Z}{2Z} \quad n \geqslant 2$$

5) $X = T^2$,二维环面.

环面与积空间 $S_1 \times S_1$ 同胚.容易建立下列公式

$$\pi_1(X \times X') = \pi_1(X) \times \pi_1(X')$$

由此推出

$$\pi_1(T^2) = Z \times Z$$

这是有两个生成元的自由阿贝尔群.

也可利用下述事实推出所得结果:平面 $\mathbf{R}^2$ 是 T^2 的万用覆叠,自同构群平移 $m\omega + m'\omega'$ 构成的群,这里 ω,ω' 是无关向量(参看椭圆函数).

可以取纬圆与子午圆为 $\pi_1(T^2)$ 的生成元 .

6) X 是"双孔"定向曲面.

在所有上述例子中,我们可以明确地给出万用覆叠 $\tilde{X} \to X$,从而用它得到 $\pi_1(X)$.对于本例这就困难得多(有时借助于自守函数可以做到).因此有必要使用别的方法.

我们把曲面 X 表为一个八边形,这八条边分别记为 $a,b,a^{-1},b^{-1},c,d,c^{-1},d^{-1}$(图5).每一条边代表 X 上的一条闭路,其同伦类分别记为 $\alpha,\beta,\alpha^{-1},\beta^{-1}$,等.显然,闭路积 $\alpha\beta\alpha^{-1}\beta^{-1}\gamma\delta\gamma^{-1}\delta^{-1}$ 与 0 同伦.于是使用类似于3)用过的那些形变法可证,$\alpha,\beta,\gamma,\delta$ 生成 $\pi_1(X)$,它们所满足的关系都是下述关系的推论

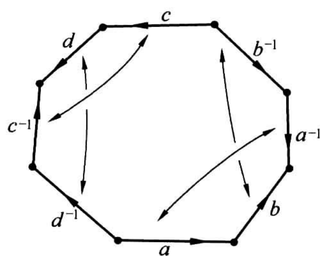


图5

$$\alpha\beta\alpha^{-1}\beta^{-1}\gamma\delta\gamma^{-1}\delta^{-1} = 1$$

因此我们有下列简写表达式

$$\pi_1(X) = \frac{\{\alpha,\beta,\gamma,\delta\}}{\alpha\beta\alpha^{-1}\beta^{-1}\gamma\delta\gamma^{-1}\delta^{-1}}$$

这是一个基本群不是阿贝尔群的空间的例子.

7 单连通概念的应用.单值化定理

单连通空间的概念与复变函数论的关系如下:

令 X 是复平面 $\mathbf{C}$ 的区域,x_0 是 X 的一点,f_0 是 $(x - x_0)$ 的幂级数,在 x_0 的一个邻域内收敛.若 x_t 是 X 上的一条道路,以 x_0 为起点,则可定义 f_0 沿着 x_t 解析开拓的概念;这是一种运算,一般说来,并不总是可能的.对于给定的级数 f_0,

我们假设开拓总是可能的,并设 X 是单连通的. 于是存在 X 上的全纯函数f,在 x_0 的一个邻域内与 f_0 重合:这就是单值化定理,意思是说,解析开拓与道路的选取无关. 对于两条相邻的道路,这是明显的;单连通的假设则可以使一条道路形变为一些依次相邻的道路而过渡到另一条道路.

当然,这个定理并非单复变函数论特有的结果,对于多复变函数,调和函数等也是对的. 唯一的条件是要求开拓的局部唯一性.

下述结果可以归结为单值化原理:

G,H 是两个连通拓扑群,f 是把 G 映入 H 的"局部同态映射",即是一个连续映射,把 G 的单位元素的一个邻域映入H,满足 $f(xy)=f(x)f(y)$,这里 x 与 y 同单位元素充分接近,于是,若 G 是单连通的,则存在把 G 映入H 的整体同态映射,它是 f 的开拓.

例子 $G=\mathbf{R},H=S_1;G=S_3,H=SO_3(3)$.

8 基本群概念的拓扑应用

群 $\pi_1(X)$ 显然是空间 X 的不变量:两个空间 X 与 X' 仅当它们的基本群同构时才可能同胚. 我们可以问:$\pi_1(X)$ 是否与 X 的其他拓扑不变量有关,特别是否与 X 的同调群有关. 1 维同调群 $H_1(X)$ 的定义方法实际上与定义 $\pi_1(X)$ 的方法很接近:闭路不过就是 1 维循环,同伦于 0 的闭路是边缘,然而其逆不真(见 L · 许瓦兹的讲演),群 $\pi_1(X)$ 与 $H_1(X)$ 一般是不同的;我们仅有:

若 $\pi_1(X)$ 是阿贝尔群,则 $H_1(X)$ 同构于 $\pi_1(X)$.

由此可知,$\pi_1(X)$ 是比 $H_1(X)$ 更精确的不变量,下面的例子属于普昂卡雷:

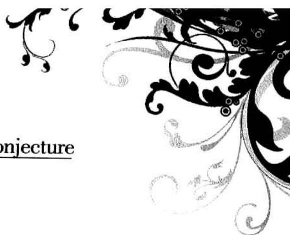
设 K 是空间 $\mathbf{R}^3$ 的旋转群,它保持正二十面体不变;已知:K 同构于五个文字的交代群A_5. 对于例 c) 的覆叠 $p:S_3\rightarrow SO(3)$,K 的逆象G是S_3的子群,S_3有 120 个元素;此外,群 G 如果是阿贝尔群,就等于 0. 于是,定义 $X=\frac{S_3}{G}$,即是 S_3 关于 G 所确定的等价关系的商群;空间 X 是 3 维紧致流形,以 S_3 为其万用覆叠,以 G为基本群. 由此可见 $H_1(x)=0$,x 与S_3有相同的同调群,虽然 $\pi_1(X)=G$ 与$\pi_1(S_3)=0$ 不同. 空间 X 称为普昂卡雷空间. 这里正好提一下普昂卡雷猜测(一直未被证明):

任何单连通的 3 维紧致流形均同胚于球面 S_3.

作为基本群在拓扑学中的另一应用,我们指出下面的纽结定理:

纽结是指 $\mathbf{R}^3$ 中与圆周 S_1 同胚的子空间 N;换言之,这是 $\mathbf{R}^3$ 中一条无重点的闭路. 两条纽结 N 与N' 叫做同构的,如果存在把整个 $\mathbf{R}^3$ 映成自身的同胚映





射，把 N 变成N'，并保持 $\mathbf{R}^3$ 的定向，同构于平面上圆周的纽结叫简单纽结(实际上没有"打结"). 怎样证明一条给定的纽结，例如"三叶纽结"(图 6)，不是简单纽结呢?注意，若 N 与 N' 同构，则它们的补集 $\mathbf{R}^3 - N$ 与 $\mathbf{R}^3 - N'$ 同胚，故有相同的基本群，这就导致研究 $\pi_1(\mathbf{R}^3 - N)$，称为纽结 N 的群；两个纽结的群如果不同构，这两条纽结就不同构. 有一种方法，可以由生成元和其间的关系来确定投射给出的纽结群. 对于简单纽结 S_1，这个群是 Z；对于三叶纽结，这个群由两个生成元 x，y 产生，其间的关系是 $xyx = yxy$. 因为这个群不是阿贝尔群(事实上，它以三个文字的对称群作为商群)，故不与 Z 同构，三叶纽结的确是有结的，正如我们所预料那样!

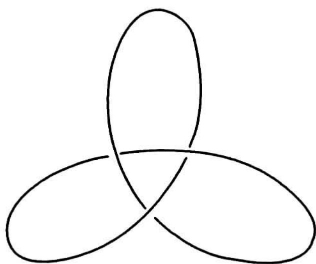


图 6

注意，这里同调群仍然不会起什么作用. 事实上，对于任何纽结 N，$H_1(\mathbf{R}^3 - N) = Z$.

9　对代数函数存在定理的应用

设 c 是复数体 $\mathbf{C}$ 上的一条代数曲线，无奇点，f 是 c 上的有理函数. 可以将 f 视为把 c 映入球面 S_2 的连续映射(将 S_2 等同于添加了无穷远点的复平面 $\mathbf{C}$). 我们不考虑 f 为常数的平凡情形. 映射 $f: c \to S_2$ 使 c 成为 S_2 的一个"多分枝覆叠"，这表明，如果从 S_2 中挖掉使映射 f 的微分为零的点，便得一个真正的覆叠 $Y \to X$，空间 X 是挖掉一些点 $P_1, \cdots, P_r$ 的 S_2，Y 是挖掉了诸 P_i 的逆象的 c. 这就是单变量代数函数的黎曼存在定理，故有结论：

挖掉了点 $P_1, \cdots, P_r$ 的球面，它的任何有限多叶的连通覆叠结构均可由上述方法唯一得到.

(唯一性容易；存在性涉及某些强有力的分析方法，第一个(正确的) 证明属于希尔伯特)

$X = S_2 - \{P_1\} - \cdots - \{P_r\}$ 的基本群是有 $r - 1$ 个生成元的自由(非阿贝尔) 群，上述结论可以化为代数函数的存在定理，这种代数函数具有若干"确定条件"，按照 $r - 1$ 个已知代替法则互相替换.

作为出发点，如果不用曲线 c 而用曲面 S，就应该把射影直线 S_2 换成射影复平面 $P_2(\mathbf{C})$，把点 $P_1, \cdots, P_r$ 换成平面代数曲线 $D_1, \cdots, D_r$. 存在(唯一) 定理仍有效；这就是恩里克定理.

第十七讲　代数拓扑:同调论初步[①]

我们先要讲一些事实,涉及普昂卡雷时代所谓的"形势分析"(l'Analysis situs)[②].因此,我们要提出"代数拓扑"这门学科中一些最基本的直观概念,其中重点讨论那些与"同调论"有关的问题.

1　直观几何的几个问题

1) 球面是三维空间的曲面,球面上有闭曲线的概念.通过保持闭曲线特征的连续形变,我们可以把闭曲线缩成一点(后面提到这个性质时就说闭曲线同伦于零).

在环面上除了上述类型的曲线 C,即是可以缩成一点的闭曲线,特别是环面上一个充分小区域内的所有闭曲线外,环面上显然还有另外一些曲线 C',例如经圆,绕中心圆周旋转,以及曲线 C'',例如纬圆,绕环面的轴旋转,曲线 C' 和 C'' 都不可能由使之保持闭合的连续形变缩成一点,也不可能彼此互变.

上面提出的问题就是同调论的研究对象.

2) 缩成一点的问题与交截问题有关:例如环面上的曲线 C' 与 C'' 必定至少相交于一点(我们不妨回想二次曲面上的两族直母线).

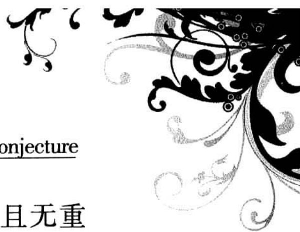
3) 区域问题.

考虑"有洞的"三维空间 $\mathbf{R}^3$,即挖去一点,例如挖去原点后的 $\mathbf{R}^3$(所以是 $\mathbf{R}^3-O$),其中每条闭曲线可缩成一点,但不是所有的闭曲面都能缩成一点:如果 O 属于一个球面的内部区域,则该球面在 $\mathbf{R}^3$ 中不能缩成一点,否则形变时将通过 O.这就提出曲面同伦的问题.我们又碰到了一张曲面(例如上述球面)与一条曲线(例如半直线 Ox)的相交问题.

如果没有精确的定义,显然不能从数学上来研究这些问题.例如,紧致曲线是指区间[0,1]的连续象,如果0与1有同样的象,就说该曲线是闭的.对于平面,我们知道约当定理的结论:任何无重点的闭曲线将平面分成两个区域.(依定义,同一区域内的两点可以用一条与给定的闭曲线无交点的连续曲线联结起来,但不同区域的两点,则不可能这样联结起来)因此发生一个交截问题,这个

① 原作者 L. Schwartz(索尔本大学教授).
② 今称拓扑学.





结果在环面上不成立,因为上面提到的曲线 C' 与 C'' 虽然都是闭曲线且无重点,但却不能把环面分成两个区域.如果考虑到这个情况,就不能认为上述结果是显而易见的,无足轻重的.

4) 向量场问题

在一个圆周上,显然可以定义一个向量场,向量与圆周相切,随圆周的点连续地变动,决不为零.然而,在 $\mathbf{R}^3$ 的一个球面上这是可能的吗?回答是“不可能”.不过,在 $\mathbf{R}^4$ 的三维球面上却又是可能的.可以证明:在 $\mathbf{R}^{n+1}$ 的球面 S^n 上,存在这样的场的充要条件是:n 为奇数.这个问题对曲面上微分方程的研究极重要,诸如存在性,奇点分类,颈状区(Col),纽结等(勒夫西兹的工作).

5) 变换的不动点存在问题具有同样性质.例如,可以证明,每个把球体映入自身的连续映射至少有一个不动点.

6) 当我们讨论复平面内留数的计算时,要研究一个曲线积分,计算积分曲线“环绕”原点或奇点多少次.这是一个正数,零或负数,依赖于所考虑的奇点同复平面无穷远点相联的一条曲线与积分曲线交点的“代数”个数.这个“环绕数”应予定义,虽然就简单情形而言在直观上是显然的.

2 曲面的边缘

我们来更详细研究一个问题,它与上述诸问题都有关,而且特别有代表性.我们只讨论紧致空间.

在一个相当简单的拓扑空间上,假设有一条闭曲线(不论是否有重点),该曲线能否是一张曲面的边缘?

首先直观地看一下上面所提问题的意思.假设(图 1)一条曲线围成一张定向曲面①,曲线的定向应该与所选定的方向一致.反之,设有一闭曲线沿确定方向绕行,它能否成为一张曲面的“边缘”?在球面上是对的,它是两张曲面的边缘,这两张曲面必有不同定向.然而在环面上呢?对于类似于经圆的曲线 C',结论不对,因为对环面定了向后,曲线 C' 上就有一个二重方向;这条曲线沿相反的方向绕行两次,所以应该看成是零.对于曲线 C'',例如纬圆,情况也是一样.因此,在环面上有些闭曲线不能成为曲面的边缘.$\mathbf{R}^2-O$ 的情形也一样,如果照我们前面所说,只考虑紧致曲面,即是不考虑无穷远点.在 $\mathbf{R}^3-O$ 内,一条闭曲线是曲面的边缘,但一张含有 O 的球面则不是边缘.

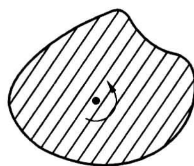


图 1

---

① 原文为在其上每一点都定向的曲面.曲面可以有“在一点定向”的概念,而且与整体定向的概念一致.但这里似与局部定向无关,故略去“在其上每一点”.

$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

上面所举例子对于“缩成一点”与“曲面的边缘”这样两个问题情况都一样，这表明这两个概念之间是有联系的．首先，缩成一点是同伦问题；其次，曲面边缘的研究是同调问题．这两问题并不等价；同伦的观点更精细，即是说，凡是可以由连续形变缩成一点的闭曲线都是曲面的边缘，凡是可缩成一点的闭曲面都是一立体的边缘，等，但其逆不确：有些曲面（立体等）的边缘，不可能由连续形变缩形一点，下面举些例子来说明：

第1个例子　在平面 $\mathbf{R}^2$ 内，挖掉两点 A 与 B，画出曲线1234，如图2所示．这曲线显然是曲面的边缘．假设按弧段1与3逆向，2与4逆向定义方向，则在同调的意义下，曲线归结为零，但是在同伦的意义下，它不能缩成一点．

第2个例子　图3表示曲面上打了两个洞，把一根管子的两头接在洞口．环绕这两个洞的闭曲线 C 是一张曲面的边缘，但是，它显然不能化成零，即使管子化成一条联结两点的线．

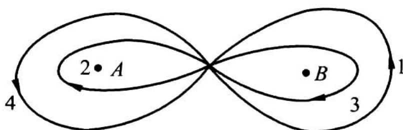
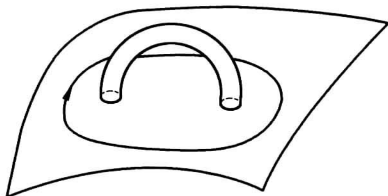


图2


图3

因此，同调论比同伦论简单．后者肇始于普昂卡雷，但人们很快就去搞同调论了．

特别，对于上面提到的留数定理，真正起作用的是同调论的观点．再提一下那个定理，开头说得很含糊：若 $f(z)$ 是平面上的一致全纯函数，则 $\int_C f(z)\mathrm{d}z = 0$，其中，$C$ 是任意闭曲线．但是，如果函数有极点，柯西曾对每一极点定义了函数在该点处的所谓留数，而柯西基本定理可以表为下述等式

$$\frac{1}{2\pi i}\int_C f(z)\mathrm{d}z = \sum R_i$$

上式右端是曲线 C“包围”的诸极点处留数的和．

就图2来说，显然可见，A 与 B 若为极点，则积分为零，因为以所说曲线为边界的曲面不含极点．但积分曲线不能（不穿过极点）缩成一点．就图4来说，积分曲线的走向如图所示，有必要规定 A 那种点的系数为1，B 那样的点的系数为2．同样，就图5及5′而言，必须使用 $+1$ 与 $+1$，或 $+1$ 与 -1．

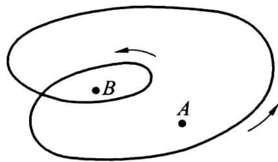
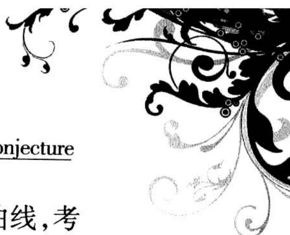


图4



因此,为了赋予一张曲面在极点邻近的系数,必须考虑曲面的边缘曲线,考虑曲线的定向以及曲面极点的阶数.

复变数的有理分式可视为黎曼球面上的亚纯函数;而该球面上的简单闭曲线 C 是两张曲面 S_1 与 S_2 的边缘,S_1 与 S_2 的并集构成整个球面,应分别赋予系数 $+1$ 与 -1,因此沿 C 的积分值等于 S_1 内诸极点的留数和减去 S_2 内诸极点的留数和.由此推出留数总和是零.然而,很明显,实际上所考虑的曲线 C 并不影响这结果;只需考虑整个球面,其边缘为零,便可得到结论.

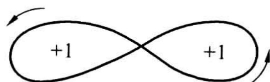


图 5

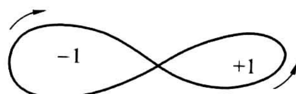


图 5′

3 确切的定义

我们刚才指出了问题是怎样产生的,使人感到了它的重要性.现在的问题是要给以精确的形式.任何数学理论都表达一种直观的思想,但应当不断使直觉升华,因为直觉有碍于论证,并有陷入"偷送论据"①的危险.我们将只讨论紧致拓扑空间,例如球面 S^2 或环面 T^2.所考虑的曲面都假设是可三角剖分的,即是可用一些单形铺满.

复形 我们已经知道(见 H·嘉当的讲演)如何在 $\mathbf{R}^{n+1}$ 中定义 n 维单形($n=0$ 时是一点;$n=1$ 时是线段;$n=2$ 时是三角形面;$n=3$ 时是四面体,等),以及如何定义单形的面,即是维数小于 n 的单形.我们说,曲面 S^2 或 T^2,或者任何紧致拓扑空间可以用单形铺满,如果下列条件满足:

1)曲面上的单形是从 $\mathbf{R}^n$ 的单形出发,经同胚映射确定的:曲面上选出的每一单形 σ 都经过一个确定的同胚映射对应于 $\mathbf{R}^n$ 的一个单形 σ^1(从而 σ 的所有维数 $\leqslant n-1$ 的面完全确定),σ 叫拓扑单形.

2)曲面被有限个单形覆盖(紧致性假设).

3)两个单形的交或者是空集,或者正好是一公共面.

如果一个空间存在这样的填铺方式,就称为复形.

一般,根据条件 2),对于这种剖分给出的复形,可以定义它的拓扑维数是所含单形的最大维数 n.但必须指出,拓扑维数可能不是处处一样(例如,考虑添上一条细线的球面).

假设曲面可三角剖分,就可以使问题大大简化;例如,一条曲线就是一维单

① 即证明中以本身尚待证明的判断作为论据的一种逻辑错误.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

形的并集.但是,每一拓扑流形都可三角剖分吗?这个问题一般没有解决.不过,提出这个简化假设是有道理的,至少开始是这样.另一方面,我们指出,一个空间即使有若干不同的三角剖分,但下面要定义的同调群却与所选取的剖分无关.

复形的链 首先定义 p 维胞腔.

p 维胞腔是排成一定次序的一组 $p+1$ 个相同或不同的点,记为 $a_0a_1\cdots a_p$.这些点是复形的同一个单形的所有顶点.

0 维胞腔是三角剖分的任一顶点 a_0.

1 维胞腔是一维单形的两个顶点的组 a_0a_1.必须注意,a_1a_0 是另一个 1 维胞腔,它与 a_0a_1 是完全不同的,a_0a_1 也视为 1 维胞腔,它的两个顶点重合.

p 取定后,p 维胞腔的个数 N_p 有限(假设 2),可以记成

$$\gamma_p^1,\gamma_p^2,\cdots,\gamma_p^{N_p}$$

p 维链就定义成 p 维胞腔的整系数的形式线性组合:

$$\gamma=\sum_{i=1}^{N_p}\xi_i\gamma_p^i\quad \xi_i\in\mathbf{Z}$$

定义链的加法法则如下

$$\left(\sum\xi_i\gamma_p^i\right)+\left(\sum\eta_i\gamma_p^i\right)=\sum(\xi_i+\eta_i)\gamma_p^i$$

这加法是结合的,交换的;零链(对于所有 $i,\xi_i=0$)是加法零元素.

因此,p 维链的集合是一个阿贝尔群,记成 Γ_p.

p 维链的边缘 这是一个 $(p-1)$ 维链,下面给出定义.一个 p 维链依定义可以写成

$$\sum_{i\in I}\xi_i\gamma_p^i$$

I 是集合 $\{1,2,\cdots,N_p\}$.

首先定义

$$\left(\sum\xi_i\gamma_p^i\right)\text{ 的边缘 }=\sum\xi_i(\gamma_p^i\text{ 的边缘})$$

这归结为定义胞腔的边缘.

a_0 是 0 维胞腔.它的边缘是零.

ab 是一维胞腔.它的边缘是 0 维链,定义成:(ab) 的边缘 $=+b-a$,记成 $\mathscr{B}(ab)$.

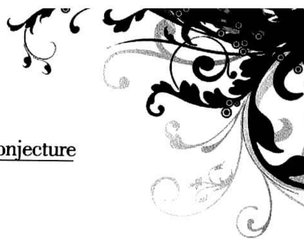
2 维胞腔 abc 的边缘是

$$\mathscr{B}(abc)=bc-ca+ab$$

一般,p 维胞腔 $a_0a_1a_2\cdots a_p$ 的边缘是 $(p-1)$ 维链

$$a_1a_2\cdots a_p-a_0a_2\cdots a_p+\cdots+(-1)^ka_0a_1\cdots a_{k-1}a_{k+1}\cdots a_p+\cdots+$$





$(-1)^p a_0 a_1 a_2 \cdots a_{p-1}$

很简单的形式计算使我们能确定 p 维链的边缘.

p 维循环(拓扑等价于 p 维闭流形)　p 维循环定义成边缘是零的 p 维链.

把 p 维链变成它的边缘的运算满足 $\mathscr{B}(\gamma_1+\gamma_2)=\mathscr{B}(\gamma_1)+\mathscr{B}(\gamma_2)$(根据边缘运算的定义),所以 p 维循环的集合 Z_p 是阿贝尔群,含于 p 维链群 Γ_p 内.

问题　在 p 维链中,$p+1$ 维链的边缘有什么性质①?

1) 边缘是一个循环,换言之,边缘的边缘恒为零.只需对一个胞腔证明 $\mathscr{B}\mathscr{B}(\gamma_p^i)=0$ 即可.例如

$$\mathscr{B}(abc)=bc-ac+ab$$

$$\mathscr{B}\mathscr{B}(abc)=c-b-c+a+b-a=0$$

推广至一般情形毫无困难.

因此,令 B_p 表示成为边缘的那些 p 维链的集合,则有包含关系:$\Gamma_p\supset Z_p\supset B_p$.

2) 这样一来,边缘群是循环群的子群.但是,这些群是否相同呢?凭直观,对于球面 S^2 答案是肯定的,而对环面 T^2 则不然.研究商群 $Z_p/B_p=H_p$ 便可作确切的比较.这商群称为所研究流形的 p 维同调群②,它刻画了所研究的结构.成为边缘的 p 维链即 B_p 的元素叫做同调于零的,记成 ~ 0.两个循环 Γ 与 Γ' 叫做同调的,如果它们属于同一个等价类,即它们的差同调于零(换言之,这差是边缘),写成 $\Gamma\sim\Gamma'$.

同调群的计算是按代数拓扑的方法进行的.我们只指出几点注意:

(1) 退化胞腔 aa 是循环,因它的边缘是 $a-a=0$.但 aa 是边缘吗?是的,它是 aaa 的边缘,故它同调于零,就是说可以把 aa 从链中消去:$aa\sim 0$.

(2) 链 $ab+ba$ 同调于零:$ab+ba\sim 0$.事实上,bab 的边缘 $\mathscr{B}(bab)=ab-bb+ba\sim ab+ba$.因此,从同调的观点看,可用 $-ba$ 代替 ab:$ab\sim -ba$.

(3) 更一般,在一个胞腔中,对顶点施行置换得到新的胞腔,与原来的胞腔是否同调,就看所用的置换是偶置换还是奇置换而定.

计算同调群的例子

1) 球面 S^2 假定已经三角剖分,0 维循环就是 0 维链 $\sum\xi_i a_i$,诸 a_i 是三角剖分的顶点.

0 维边缘是 0 维循环 $\sum\xi_i a_i$,使得 $\sum\xi_i=0$.这个条件显然是必要的,因为它对于胞腔的边缘为真.它又是充分的,因为在球面上,可以把一个选定的顶点 a_0 与任何顶点 a_i 联结起来,所以链 $(a_i-a_0)\sim 0$.然而,若 $\sum\xi_i=0$,则

① 这句的原文是:p 维循环中,$(p-1)$ 维链的边缘有什么性质?

② 提醒一下,若 Z 是用加法表示的阿贝尔群,B 是其子群,则 $H=Z/B$ 是由下列关系确定的等价类的集合:$x\in Z,y\in Z,x\equiv y\Leftrightarrow x-y\in B$.

$\sum \xi_i a_i = \sum \xi_i (a_i - a_0) \sim 0.$

于是,两个0维循环等价即同调,如果它们有同样的整数值$\sum \xi_i$.这表明同调群 $H_0 = Z_0/B_0$ 是整数群.

我们看到只要任意两点都可以用 1 维单形构成的折线联结起来,就是说,所考虑的复形是连通拓扑空间,那么上述结论仍然成立.例如,不相交的两个球面的并集就不是这样,这时,H_0 是整数对所成的群,它给出了连通区的个数.

现在研究 $H_1 = Z_1/B_1$.可证在球面上每个 1 维循环都是边缘,故 B_1 等于 Z_1,因此 H_1 是仅有零元素 0 所成的群(为了完成证明,要利用三角剖分,例如,分成 8 个三角形,它们的边都是四分之一的大圆).

我们不讨论二维同调群了.

2) 在一点 0 处穿孔的平面,即是 $\mathbf{R}^2 - 0$.这不是紧致空间,但我们规定只使用有限多个胞腔组成的链,以便继续使用有限同调论或具有紧致支集的同调论的方法.可以证明,能够定义一个基本循环 γ_1,它的同调类用 1 表示,使得任何循环 γ 都同调于基本循环的n 倍(从而 $\gamma - n\gamma_1 \sim 0$,这是边缘),这里整数 n 的符号任意.按定义,相应于整数 n 的同调类,就是“围绕点 0 的代数次数为 n 的”那些循环组成的等价类.因此,1 维同调群是整数群.

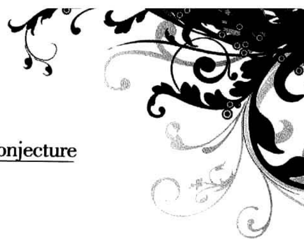
3) 环面 T^2.用代数拓扑的方法可以证明:环面的同调群是整数对(m,n)所成的群,其加法法则为

$$(m,n) + (m',n') = (m + m',n + n')$$

如果我们用$(1,0)$代表同调于环面经圆的曲线 C',用$(0,1)$代表同调于环面纬圆的曲线 C'',则每一循环同调于 $mC' + nC''$,m 与 n 是唯一确定的.因此,我们说,在代数意义下,这个循环围绕中心圆周 m 次,围绕旋转轴 n 次.例如 Villarceau 圆环绕旋转轴及中心圆各一次,其同调类是$(\pm 1, \pm 1)$①.

① C' 与 C'' 都有确定的走向,m 与 n 是符号任意的整数.有两族 Villarceau 圆,每一族都有两个走向.因此给出四个同调类:$(1,1)$,$(1,-1)$,$(-1,1)$,$(-1,-1)$.





附录Ⅳ Manifolds with Density and Perelman's Proof of the Poincaré Conjecture①

Perelman's 2003 proof of the 1904 Poincaré Conjecture considers a manifold with a density (as in freshman physics or calculus), as we will explain. Manifolds with density long have appeared in mathematics, with more recent attention to their differential geometry, including a generalization of Ricci curvature, which Perelman uses in exploring the Ricci flow. There are useful isoperimetric results, such as a generalization of the Levy – Gromov isoperimetric comparison theorem. In the 1980s, Bakry and Ledoux used Markov semigroup arguments. Bayle used first and second variation formulas. The approach here, following and correcting our survey via Heintze – Karcher, seems to be the simplest. The goal is to generalize all of Riemannian geometry to manifolds with density.

1 Definitions A *manifold with density* is a Riemannian manifold M^n with a positive density function $\Psi(x)$ used to weight volume and hyperarea (and sometimes lower-dimensional area and length). In terms of the underlying Riemannian volume $\mathrm{d}V_0$ and area $\mathrm{d}A_0$, the new, Weighted volume and area are given by

$$\mathrm{d}V = \psi \mathrm{d}V_0, \mathrm{d}A = \psi \mathrm{d}A_0$$

Such a density is not equivalent to scaling the metric conformally by a factor $\lambda(x)$ since in that case volume and area would scale by different powers of λ. Manifolds with density, a special case of the "mm spaces" of Gromov or the earlier "spaces of homogeneous type" (see [Coifman and Weiss, pp. 587, 591]), long have arisen on an *ad hoc* basis in mathematics. Quotients of Riemannian manifolds are manifolds with density. For example, R^3 modulo rotation about the z-axis is the half-plane

$$H = \{(x, z) : x \geqslant 0\}$$

Figure 1. Gauss space is Euclidean space with Gaussian density, which drops off exponentially. Image by Diana Davis.

with density $2\pi x$; volume and area in R^3 are given by integrating this density over the generatinr region or curves in H. A manifold with density of much interest to probabilists is *Gauss space* G^n (see Ledoux-Talagrand, McKean, or Stroock), which is Euclidean space with Gaussian density

① 转载自 Jams F. Bredt. Frank Morgan.《几何测度论》(第4版),世界图书出版公司,2009.

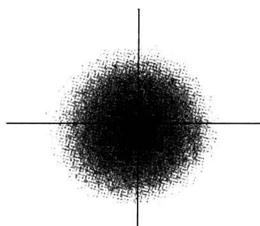


Figure 1

$$\Phi = (2\pi)^{-n/2} e^{-x^2/2} \tag{1}$$

as in Figure 1. The leading coefficient just gives G^n unit weighted volume. Note that Gauss space is a product space and is rotationally invariant. Any such density on R^n must be a multiple of e^{cx^2}. The normalized scalings of Gauss space have density $(\gamma/2\pi)^{n/2} e^{-\gamma x^2/2}$.

In Gauss space, the isoperimetric problem has an even nicer solution than the hyperspheres of Euclidean space or spheres. Because of the rapidly decreasing density, a perimeter-minimizing surface could well be unbounded, and it is.

2 Theorem (Sudakov-Tsirel'son, Borell) *In Gauss space, hyperplanes minimize area for given volume.*

Proof Sketch Mehler observed that for fixed n, Gauss space G^n can be obtained as the limit of orthogonal projections into R^n of high-dimensional spheres $\mathbf{S}^{N-1}$ of radius $\sqrt{N}$, with density normalized to yield unit volume. To see this, first make the trivial observation that since G^N is a product space, G^n is the orthogonal projection of G^N for $N \geqslant n$. The unrelated components of a vector in G^N have Gaussian distributions (in other words, they are independent Gaussian random variables). An easy computation shows that the square of each component has expected value 1. By the central limit theorem, the sum of the squares is probably close to N; that is, the measure is (uniformly) concentrated on the sphere of radius $\sqrt{N}$. Thus, G^n is the limit of projections of such spheres.

The inverse projection of a hyperplane in G^n is a hypersphere in S^{N-1}. Since hyperspheres are known to be area minimizing for given volume in S^{N-1}, a careful limit arguments shows that hyperplanes are area minimizing in G^n.

Theorem 7 and Corollary 10 provide two other proofs that also show uniqueness, first proved by Carlen and Kerce. The fact that not only Gauss space but also its isoperimetric regions are products makes it an excelent model with implications for other spaces. For example, it happens to follow quickly from Theorem 2 that the least-perime-



ter way to divide a cube in half is a wall down the middle, a result first proved by Hadwiger by a clever, *ad hoc* argument.

For applications to Brownian motion and to stock option pricing.

18.3 Curvature There is a canonical extension to smooth manifolds with (smooth) density e^{ψ} of the notion of the mean curvature of a smooth hypersurface: as in the classical case, it is the quantity that appears in the first variation formula for how the area changes as the surface moves. The formula for the generalized mean curvature H_{ψ} involves both the classical mean curvature H and the log of the density

$$H_{\psi} = H - \frac{1}{n-1}\frac{\mathrm{d}\psi}{\mathrm{d}n} \tag{1}$$

General regularity for minimizers extends to manifolds with density. In particular, for a perimeter-minimizing surface for fixed volume, H_{ψ} must be constant since otherwise, by the first variation formula, moving the surface outward where H_{ψ} is larger and inward where H_{ψ} is smaller could preserve volume and decrease perimeter. As in classical differential geometry, $-(n-1)H_{\psi}$ has the interpretation $\mathrm{d}P/\mathrm{d}V$.

For a two-dimensional smooth manifold with density e^{ψ}, Corwin *et* al. define a generalized Gauss curvature

$$G_{\psi} = G - \Delta\psi \tag{2}$$

and obtain a generalization of the Gauss-Bonnet formula for a smooth disc R:

$$\int_R G_{\psi} + \int_{\partial R} \kappa\psi = 2\pi \tag{3}$$

where k_{ψ} is the inward one-dimensional generalized mean curvature as in(1) and the integrals are with respect to unweighted Riemannian area and arclength. Certain punctured planes with density with $G_{\psi} = 0$ were studied by Carroll *et al*. Different generalization of Gauss curvature, involving $|\nabla\psi|^2$, are needed to recover asymptotic formulas for areas and perimeters of small discs.

In general dimensions, there are various useful generalizations of Ricci curvature (see Bayle, Villani, and references therein to Bakry, Émry, Ledoux, and others; also Chang et al.), generally involving Hess ψ and $\mathrm{d}\psi \otimes \mathrm{d}\psi$. My favorite generalization of Ricci curvature is simply

$$\mathrm{Ric}_{\psi} = \mathrm{Ric} - \mathrm{Hess}\ \psi \tag{4}$$

the generalized curvature of Lichnerowicz, Bakry – Émery[Prop. 3], and Bakry-Ledoux [p.265]. It yields the simplest generalization of Levy-Gromov 7, appears in the second variation formula 9, and appears in Perelman's paper on the Poincaré Conjecture. For Gauss space $G^n(1)$ with density $\Phi = \mathrm{e}^{\varphi}$ it is constant

$$\mathrm{Ric}_{\varphi} = 0 - \mathrm{Hess}\ \varphi = I$$

Unlike the classical case, for a two-dimensional manifold with density, the Gauss curvature(2) is not quite half the trace of the Ricci curvature(4).

The classical Heintze-Karcher formula(see also[Burago-Zalgaller, 34.1.10(11)]) provides an upper bound on the volume of a one-sided neighborhood of a hypersurface in terms of its mean curvature and the Ricci curvature of the ambient manifold.

4 Theorem(Classical Heintze-Karcher) *Let M^n be a smooth, complete Riemannian manifold with Ricci curvature $Ric \geqslant (n-1)\delta$. Let S be a smooth, oriented, closed hypersurface in M with mean curvature $H(s)$. Let $V(r)$ denote the volume of the region within distance r of S on the side of the unit normal(which determines the sign of H). Then*

$$V(r) = \int_S \int_0^{r^*(s)} [c_\delta(t) - H(s)s_\delta(t)]^{n-1} \mathrm{d}t\mathrm{d}s \tag{1}$$

where ds denotes weighted surface area

$$S_\delta(t) = \begin{cases} \delta^{-1/2}\sin \delta^{1/2} t & \text{for } \delta > 0 \\ t & \text{for } \delta = 0 \\ |\delta|^{-1/2}\sinh|\delta|^{1/2} t & \text{for } \delta < 0 \end{cases}$$

$c_\delta(t) = \mathrm{d}s_\delta(t)/\mathrm{d}t$, and r^* is the lesser of r and the first zero of $c_\delta(t) - H(s)s_\delta(t)$.

The result generalizes to closed surfaces of higher codimension.

The following generalization of Heinze-Karcher to manifolds with density is simpler and has an easier proof.

5 Theorem (Heintze-Karcher for Manifolds with Density) *Let M^n be a smooth, complete Riemannian manifold with smooth density $\psi = e^{\varphi}$ satisfying*

$$\mathrm{Ric}_{\psi} = \mathrm{Ric} - \mathrm{Hess}\ \psi \geqslant \gamma$$

Let S be a smooth, oriented, complete, finite-area hypersurface in M with generalized mean curvature

$$H_\psi(s) = H(s) - \frac{1}{n-1}\frac{\mathrm{d}\psi}{\mathrm{d}n}$$

Let $V(r)$ denote the volume of the region within distance r of S on the side of the unit normal(which determines the sign of H_ψ). Then

$$V(r) \leqslant \int_S \int_0^r \exp(-(n-1)H_\psi(s)t - \gamma t^2/2)\mathrm{d}t\mathrm{d}s \tag{1}$$

where ds denotes weighted surface area.

If equality holds, then S has vanishing classical second fundamental form, the region is a metric product $S \times [0, r]$, and inside the region, along geodesics normal





to S, $-\mathrm{d}^2\psi/\mathrm{d}t^2 = \gamma$.

Theorem 18.5 is sharp for hyperplanes in Gauss space, for example.

Proof We begin with the case $\psi = 1$. This case follows from Theorem 4, but we want to incorproate variable Ricci curvature, as in(3)below, and give an easier proof. Consider the volume element $e^{f(s,t)}\mathrm{d}t\mathrm{d}s$, corresponding to an infinitesimal slice $\mathrm{d}t$ of an infinitesimal normal wedge from S. Then $e^{f(s,t)}\mathrm{d}s$ represents an element of surface area parallel to S. By the first variation formula 9, for example, its derivative $f' e^f ds$ equals $-(n-1)He^f\mathrm{d}s$, so that

$$f' = -(n-1)H$$

Since $-(n-1)H' = \mathrm{II}^2 + \mathrm{Ric}(n,n)$, where II is the second fundamental from

$$f'' = -\mathrm{II}^2 - \mathrm{Ric}(n,n) \leqslant -\mathrm{Ric}(n,n)$$

with equality only if II^2 vanishes. Hence, by Taylor's theorem

$$f(s,t) \leqslant -(n-1)H(s)t - \int_0^t \mathrm{Ric}(n,n)\mathrm{d}t \tag{2}$$

Consequently, since every point of $V(r)$ is covered by the infinitesimal wedge from the nearest point of S

$$V(r) \leqslant \int_S \int_0^r \exp\left(-(n-1)H(s)t - \int_0^t \tau \mathrm{Ric}(n,n)\mathrm{d}\tau\right)\mathrm{d}t\mathrm{d}s_0 \tag{3}$$

where we now write ds_0 to emphasize that this is the case of unweighted area.

For general density $\psi = e^{\psi}$

$$\psi(s,t) \leqslant \psi(s,0) + t\frac{\mathrm{d}\psi}{\mathrm{d}n}(s,0) + \int_0^t \tau\frac{\mathrm{d}^2\psi}{\mathrm{d}n^2}\mathrm{d}\tau \tag{4}$$

Preparing to add f and ψ, note that

$$-(n-1)H(s) + t\frac{\mathrm{d}\psi}{\mathrm{d}n}(s) = -(n-1)H_{\psi}(s)$$

and that by hypothesis

$$-\int_0^t \tau\mathrm{Ric}(n,n)\mathrm{d}\tau + \int_0^t \tau\frac{\mathrm{d}^2\psi}{\mathrm{d}n^2}\mathrm{d}\tau \leqslant -\gamma\int_0^t \tau\mathrm{d}\tau = -\gamma t^2/2$$

Hence

$$f(s,t) + \psi(s,t) \leqslant \psi(s,0) - (n-1)H_{\psi}(s)t - \gamma t^2/2$$

Therefore

$$V(r) \leqslant \int_S\int_0^r e^f e^{\psi}\mathrm{d}t\mathrm{d}s_0 \leqslant$$

$$\int_S\int_0^r \exp(-(n-1)H_{\psi}(s)t - \gamma t^2/2)\mathrm{d}t(e^{\psi(s,0)}\mathrm{d}s_0)$$

as desired. If equality holds, the parallel hypersurfaces are disjoint and have $\mathrm{II} = 0$ so that the normal geodesics stay equidistant, which means that the region is a metric prod-

uct.

Remark Theorems 4 and 5 and their proofs apply to perimeter minimizers S with singularities. Indeed, for any point off S, the nearest point on S is a regular point because the tangent cone lies in a half-space and hence must be a hyperplane.

The classical isoperimetric theorem of Levy and Gromov compares a compact manifold of positive Ricci curvature to the model sphere. The statement is simplified by using non-unit constant density to normalize to unit volume.

6 Theorem(Classical Levy-Gromov) *Let M^n be a smooth, complete, connected Riemannian manifold with constant density, unit volume, and Ricci curvature $Ric \geqslant \gamma > 0$. Then the isoperimetric profile $P(V)$ (least perimeter to enclose given volume) satisfies*

$$P \geqslant P_{S\gamma}$$

where $P_{S\gamma}$ is the isoperimetric profile of the (constant-density, unit-volume) round sphere of Ricci curvature γ.

Theorem 7 gives a generalization to a manifold M with variable density, with the sphere replaced by Gauss space as the model of comparison. (Gauss space beats the sphere.) It includes the sharp isoperimetric inequality for Gauss space, Proposition 2. If M has finite volume, we may assume by scaling the density that M has unit volume. Such scaling does not affect the generalized Ricci curvature because multiplying the density e^{ψ} by a constant just adds a constant to ψ and leaves Hess ψ unchanged.

7 Theorem(Levy-Gromov for Manifolds with Density) *Let M^n be a smooth, complete, connected Riemannian manifold with smooth density $\psi = e^{\psi}$, unit volume, and generalized Ricci curvature*

$$\mathrm{Ric}_{\psi} = \mathrm{Ric} - \mathrm{Hess}\ \psi \geqslant \gamma > 0 \tag{1}$$

Then the isoperimetric profile $P(V)$ (least perimeter to enclose given volume) satisfies

$$P \geqslant P_{G\gamma} \tag{2}$$

where $P_{G\gamma}$ is the isoperimetric profile of (scaled) Gauss space with density

$$\Phi = e^{\varphi} = (\gamma/2\pi)^{n/2} e^{-\gamma x^2/2}$$

so that $\mathrm{Ric}_{\varphi} = -\mathrm{Hess}\ \varphi = \gamma$.

In Gauss space, perimeter minimizers are hyperplanes. If equality holds in (2) for some $0 < V < 1$, then M is a product of 1D Gauss space with some manifold with density.

Remark $P_{G\gamma}$ is independent of dimension, because the volume of a hyperplane bounding given volume is independent of dimentsion, because Gauss space is a product of 1D Gauss spaces. Hence, $P_{G\gamma}$ is just the value of the Gaussian density



$$\sqrt{\frac{\gamma}{2\pi}}e^{-\gamma x^2/2}$$

at the endpoint of a half-line with the given mass.

Proof Note that the generalized curvature H_φ of the hyperplane $S_0 = \{x_1 = a\}$ in Gauss space is given by

$$H_\varphi = H - \frac{1}{n-1}\frac{d\varphi}{dn} = 0 - \frac{1}{n-1}\gamma a$$

which is constant (as it must be for any minimizer). Let P be the perimeter given by such a minimizing hypersurface S in M and let P_0 be the perimeter given by the hyperplane S_0 in Gauss space. By replacing V by $1 - V$ (which changes the sign of the mean curvatures) if necessary, we may assume that the generalized mean curvature of S is greater than or equal to that of S_0. By generalized Heintze-Karcher (Theorem 4)

$$\frac{V}{P} \leqslant \frac{V}{P_0} = \int_0^\infty \exp(\gamma at - \gamma t^2/2)dt$$

Taking M to be Gauss space, we conclude that hyperplanes are perimeter minimizing, then equality holds in Theorem 5(1) for $r = \infty$ on both sides of S, S has vanishing classical second fundamental form in the underlying Euclidean space, and hence S is a hyperplane. We conclude that in Gauss space, hyperplanes are uniquely perimeter minimizing.

Returning to general M, we conclude that $P \geqslant P_{G\gamma}$. If equality holds, the parallel hypersurfaces are disjoint and totally geodesic ($\mathrm{II}^2 = 0$) so that the normal geodesics stay equidistant, which means that the region is a metric product.

The theorem of (Bonnet and) Myers states that the diameter of a smooth, connectd, complete, nD Riemannian manifold with Ricci curvature at least $a > 0$ is at most $\pi\sqrt{(n-1)/a}$. *My favorite generalization to manifolds with density* uses a different replacement for the Ricci curvature.

8 Myers' Theorem with Density *Let M^n be a smooth, connected Riemannian manifold with smooth density $e^\psi \leqslant b$, complete in the weighted metric. If (in the unweighted metric)*

$$\mathrm{Ric} - \Delta\psi + \mathrm{Hess}\ \psi \geqslant a > 0$$

then the weighted diameter satisfies

$$\mathrm{diam} \leqslant \pi b\sqrt{(n-1)/a}$$

The proof is a minor modification of the standard proof of Myers' theorem. Since the only weighted quantity to appear in the statement is length, there is (another, different) such result by applyinig the classical Myers' theorem to the conformally altered metric, but

the hypothesis in terms of the original metric becomes unwieldy. Qian provides other generalizations which depend on $|\nabla\psi|^2$. See also Villani, Wei-Wylie, and references therein. There are also various volume estimates generalizations of Bishop's theorem.

The alternative approach of Bayle, similar in essence to Bakry-Émery, uses just second variation. For the record, we present such formulas for manifolds with density. The classical formulas are augmented by terms involving the first and second normal derivatives of ψ. For applications to stability of a slightly reformulated second variation formula.

9 First and Second Variation *Let M^n be a smooth Riemannian manifold with smooth density $\psi = e^{\psi}$. Let S be a smooth hypersurface and consider a smooth normal variation of compact support of constant velocity $u(s)$ along the geodesic normal to S at s. The the first and second variations of (weighted) area satisfy*

$$\delta^1(u) = -\int_S u(n-1)H_\psi \tag{1}$$

$$\delta^2(u) = \int_S |\nabla u|^2 + u^2(n-1)^2H_\psi^2 - u^2\mathrm{II}^2 - u^2\mathrm{Ric}(n,n) + u^2(\mathrm{d}^2\psi/\mathrm{d}n^2) \tag{2}$$

where II *is the second fundamental form and the generalized mean curvature H_ψ satisfies*

$$H_\psi = H - \frac{1}{n-1}\frac{\mathrm{d}\psi}{\mathrm{d}n}$$

The integrals are with respect to weighted area.

Remark If $u=1$, then the second variational formula is equivalent to the first plus the fact that

$$(n-1)\frac{\mathrm{d}H}{\mathrm{d}t} = \mathrm{II}^2 + \mathrm{Ric}(n,n)$$

which follows from the case of the curvature κ of a curve in a surface of Gauss curvature G

$$\frac{\mathrm{d}\kappa}{\mathrm{d}t} = \kappa^2 + G$$

10 Corollary *Let M^n be a smooth, complete, finite volume Riemannian manifold with smooth density $\psi = e^{\psi}$ with*

$$\mathrm{Ric} - \mathrm{Hess}\ \psi \geqslant \gamma$$

Then the isoperimetric profile P and its derivatives P', P'' satisfy

$$PP'' \leqslant -\gamma \tag{1}$$

almost everywhere (and in a weak sense everywhere). If equality holds, then a perimeter





minimizer has vanishing classical second fundamental form.

In Gauss space, equality holds and perimeter minimizers are hyperplanes.

By scaling the density, which does not affect Hess ψ or PP'' or hence the hypothesis or conclusion of Corollary 10, one may assume that M has unit volme. Since equality holds in(1)for Gauss space, isoperimetric estimates such as Levy-Gromov (Theorem 7) follow; here, for example, we deduce the isoperimetric inequality in Gauss space.

Proof When we say that(1)holds in a weak sense, we mean that at each point, $P(V)$ lies under a smooth function with the same value and second derivative at most $-\gamma/P$. To prove this in the case that a perimeter-minimizing hypersurface is smooth, as always holds if $n \leqslant 7$, we just consider uniform normal perturbations, whose perimeters $P(V)$ give an upper bound on the least perimeter. To compute $P'(V)$ and $P''(V)$ for this new family, we use the variation formulae with $u = 1$. Using dots for time derivatives, by the first variation formula, since $-(n-1)H_\psi = P'$

$$\dot{P} = PP'$$

In the second variation formula, the first term of the integrand vanishes. The second term is P'^2. The third term is nonpositive and vanishes if and only if $\mathrm{II} = 0$. The last two terms are by hypothesis at most $-\gamma$. Therefore

$$\ddot{P} \leqslant P(P'^2 - \gamma)$$

Since $\dot{V} = P$, by the chain rule $P' = \mathrm{d}P/\mathrm{d}V = \dot{P}/P$ and

$$P'' = \frac{\ddot{P}P - \dot{P}^2}{P^3} \leqslant \frac{P^2(P'^2 - \gamma) - P^2P'^2}{P^3} = \frac{-\gamma}{P}$$

Therefore, $PP'' \leqslant -\gamma$, and equality implies that the perimeter minimizer has $\mathrm{II} = 0$, as desired.

Even if the perimeter-minimizing surface is not smooth, the singular set is very small, of codimension at least 8, and one can use limits of variations u that vanish in a small neighborhood of the singular set.

To prove minimizers in Gauss space, let $P_0(V)$ denote the perimeter of a hyperplane enclosing volume V, for which equality holds in(1) with $\gamma = 1$. Suppose that for some V_0, $P(V_0) < P_0(V_0)$ and that V_0 is chosen to maximize $P_0(V_0) - P(V_0)$. Since $P''(V_0) \leqslant -1/P(V_0) \leqslant -1/P_0(V_0) = P''_0(V_0)$, this yields an immediate contradiction. Hence, $P = P_0$, equality always holds in(1), and hyperplanes are minimizers. Moreover, any minimizer has $\mathrm{II} = 0$. Since the underlying Riemannian manifold is Euclidean, any minimizer is a hyperplane.

Remarks For unit or constant density, (1)is not sharp for the model case of the sphere. In that category, there is a sharp result

$$PP'' \leqslant -P'^2/(n-1) - \gamma \tag{2}$$

obtained by keeping the II^2 term and using the estimate $\mathrm{II}^2 \geqslant (n-1)H^2$, which does not hold in the category of manifolds with density for H_ψ.

Unfortunately, formulas (1) and (2) appear incorrectly in Morgan and Morgan and Johnson, respectively.

11 Perelman's Proof of the Poincaré Conjecture It has long been known that among connected compact two-dimensional manifolds, such as the sphere, the torus, and the two-holed torus, the sphere is characterized by the fact that any loop can be contracted to a point, whereas a loop around a torus, for example, cannot be contracted to a point. The Poincaré Conjecture, suggested by Henri Poincaré in 1904, proposes the analogous result for threedimensional manifolds: *a simply connected compact three-dimensional manifold must be a sphere*. At the 2006 International Congress of Mathematicians, Grigori Perelman was awarded the Fields Medal for its proof, although he declined to accept it.

High-dimensional versions of the Poincaré Conjecture, with more space to do geometric constructions, are easier. Stephen Smale proved the analogous conjecture for dimensions at least 5 and won the Fields Medal in 1966. Michael Freedman proved the 4-dimensional case and won the Fields medal in 1986.

Figure 2. In 1904, Henri Poincaré (left) proposed his famous conjecture. In 2003, Grigori Perelman (right) proved it. (Photos from the Internet)



Figure 2

The basic idea of Perelman's proof, due to Richard Hamilton, is to start with any simply connected compact three-manifold and let it shrink at each point in each direction at a rate proportional to its Ricci curvature. If you can show that you eventually end up with a round sphere, with perhaps other spheres pinched off along the way, you can conclude that you must have started with a (deformed) sphere.

The fundamental difficulty is to obtain some control over the formation of singularities. To focus attention about a point p of concern, Perelman gives the manifold large density about that point and lets the metric flow by the associated generalized Ricci cur-



vature Ric_{ψ} 3(4). If one fixed the measure, then the density evolves as a modified backwards heat equation and approaches a delta function at p. His general monotonicity of energy now provides the requisite local information instead of the usual global information. (Actually, in the proof, to facilitate surgery, Perelman moves to a localized version of the density called the length function)

Remarkably, modulo diffeomorphisms, this generalized Ricci flow is equivalent to the standard Ricci fow. As Perelman puts it, "The remarkable fact here is that different choices of [density] lead to the same flow, up to a diffeomorphism; that is, the choice of [density] is analogous to the choice of gauge." This means that during the proof one can choose any density for convenience. In summary, manifolds with density provide a convenient technical context for applying diffeomorphisms to focus attention on regions of concern.

In addition to general manifolds with density, the premier example of Gauss space makes an important appearance in Perelman's proof. Specifically, the isoperimetric inequality on Gauss space in functional form, known as the Gaussian logarithmic Sobolev inequality, is an important technical ingredient.

General manifolds with density have not appeared in most expositions of Perelman's work. This is partly because they appear just in his introductory sections and are replaced by his length function in the actual proof. It is also partly because manifolds with density are not as familiar to the mathematical community as they should be. They appear in the very first sentences of the body of Perelman's paper, where he begins by considering a manifold with density e^{-f} and scalar curvature R:

> Consider the functional $F = \int_M (R + |\nabla f|^2) e^{-f} dV$ for a Riemannian metric g_{ij} and a function f on a closed manifold M.

Anderson, in his *Notices* survey, makes the following comments on Perelman's discussion of Ricci flow in manifolds with density:

> It turns out that, given any initial metric $g(0)$ and $t > 0$, the [density] function f... can be freely specified at $g(t)$... Perelman then uses this freedom to probe the geometry of $g(t)$ with suitable choices of f. For instance, he shows by a very simple study of the form of [the energy] F that the collapse or noncollapse of the metric $g(t)$ near a point x_0 can be detected from the size of F by choosing f to be an approximation to a delta function centered at x. The more collapsed $g(t)$ is near x, the more negative the value of F. The collapse of the metric $g(t)$ on any scale in finite time is then ruled out by combining this with the fact that the functional F is increasing along the Ricci flow. In fact, this argument is carried out with respect to a some what more complicated *scale-invariant* functional than F; motivated by certain analogies in statistical physics, Perelman calls this the entropy functional.

Perelman, in the following excerpt from his Introduction, refers to this noncollapse result of his § 4. He also mentions the use in his § 3 of the Gaussian isoperimetric inequality to prove that you do not get periodicity instead of the desired limiting behavior.

> We prove that Ricci flow, considered as a dynamicaly system on the space of riemannian metrics modulo diffeomorphisms and scaling, has no nontrivial periodic orbits. The easy (and known) case of metrics with negative minimum of scalar curvature is treated in § 2; the other case is dealt with in § 3, using our main monotonicity formula (3.4) and the Gaussian logarithmic Sobolev inequality, due to L. Gross. In § 4 we apply our monotonicity formula to prove that for a smooth solution of a finite time interval, the injectivity radius at each point is controlled by the curvatures at nearby points. This result removes the major stumbling block in Hamilton's approach to geometrization.





附录Ⅴ 下个世纪的数学问题①

V.I.Arnold代表国际数学联盟已经给很多数学家写了信,建议他们描述一下下个世纪的大问题.这篇报告是我的回答.

Arnold发出这种邀请部分地受到了1900年Hilbert列出的23个问题的启发.我也利用了这些问题来帮助构思这篇文章.

我列出了18个问题,选题的准则如下:

(1)陈述简单,同时也注意数学上的确切性.

(2)我本人比较熟悉的问题,并且已经发现不是个容易的题目.

(3)一个信念:所提问题,它的解,部分结果甚或为求得解答所作的努力都可能对数学及其在下个世纪的发展有极大的重要性.

这些问题中有一些是人们熟知的.事实上,其中已包含了我相信是三个最大的未解决的问题:Riemann猜想、Poincaré猜想及"P = NP吗?".除了Riemann猜想外,下面还有一个问题也出现在Hilbert列出的问题表中(Hilbert第16问题).本文有一些内容与我早些时候的一篇文章有点相重.那篇文章的题目是:动力学的回顾,大问题及失败的尝试.

我们开始吧.

问题1:Riemann假设

Riemann zeta函数定义为

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^{s}},\mathrm{Re}(s)>1$$

的解析延拓.此假设是关于它的零点的:在区间带$0\leqslant \mathrm{Re}(s)\leqslant 1$中的零点都在$\mathrm{Re}(s)=\frac{1}{2}$的线上.

这是Hilbert所列的第8问题.有很多关于ζ函数和Riemann假设的好书,也很容易找到.这里我不再赘述.

问题2:Poincaré猜想

假定紧连通的3维流形具有如下性质:它的每个圆周均可形变为一个点,

① 原题:Mathematical Problems for the Next Century.作者:Steve Smale.译自:The Mathematical Intelligencer.Vol.20,No.2,1998.7-13.
这是1997年6月,Smale在多伦多,Fields研究所祝贺Arnold 60岁生日时所做的演讲.

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t}=-2R_{ij}$$

那么此流形必与3维球面同胚吗？

n 维球面指

$$\{x \in \mathbf{R}^{n+1} \mid \|x\| = 1\}, \|x\|^2 = \sum_{i=1}^{n+1} x_i^2$$

紧 n 维流形可以看成某个欧氏空间中的有界闭 n 维曲面(可以微分且无奇点).

n 维的 Poincaré 猜想是说：n 维紧流形 M 如果具有性质：当 $k < n$(或等价地，$k \leqslant \frac{n}{2}$)时，每个映射 $f: S^k \to M$ 均可变形为一点的话，那么它必与 S^n 同胚.

Henri Poincaré 在他开创拓扑学的先驱工作中研究了这些问题.Poincaré 在1900年给了一个一般的 n 维情形的证明.随后(在1904年)他发现他原先的说法有反例.在第二篇文章里，他只考虑 $n = 3$，并将3维情形作为如上问题(实际上不是"猜测")予以陈述.

我自己和这个问题的关系在[Smale,1990]中有交代，我在该文中写道：

"我是1955年在爱娜堡(Ann Arbor)写一篇有关拓扑问题的论文时第一次听说 Poincaré 猜想.不久，我以为我发现了一个证明(3维情形).Hans Samelson 在他的办公室里，我很激动地向他勾画了我的想法……出了办公室，我发现我的'证明'完全没有用到3维流形的任何条件."

1960年，我给 $n > 4$ 时的 n 维 Poincaré 猜测以一个肯定的回答.1983年，Mike Freedman 对 $n = 4$ 给出一个肯定的回答(注意：对 $n > 4$，我证明了更强的结果：M 是两个实心球的光滑并：$M = D^n \cup D^n$；至今，这个结果在 $n = 4$ 时仍没有证明).

与此有关的一些背景材料，除上面已经提到的外，可见[Smale,1963].

在 Poincaré 之后的许多别的数学家曾经宣称证明了3维的情形，对这些尝试的介绍见[Taubes,1987].

Poincaré猜想在数学史上是基本的，原因为：它很好地突出了流形是一个有其自身重要性的研究对象.Poincaré 就这样以其对几何对象，包括代数簇、黎曼流形等的关注而对20世纪的数学影响至巨.

我坚信，今天在"多项式时间的算法"上，存在着类似的现象.算法不仅仅只是作为解决其他问题的一种手段，它还由于其自身的重要性而显得重要.因此我认为，就像研究方程的解的集合(流形)在20世纪的数学里扮演了重要角色那样，找解的研究(算法)在下一个世纪里将会起到同样的重要作用.

问题3：P = NP吗？

有时我把这问题看成是计算机科学送给数学科学的礼物.也许把这问题写成传统数学的形式更为有用.

为此目的，首先考虑 Hilbert 在复数域上的零点定理(Nullstellensatz).为此





设 $f_1,\cdots,f_k$ 是 n 个变元的复多项式.要求判定:它们是否有一个公共的零点 $\zeta\in\mathbf{C}^n$?零点定理断言:不存在如此的 ζ 当且仅当有 n 个变元的复多项式 $g_1,\cdots,g_k$ 满足下述多项式的恒等式

$$\sum_{i=1}^{k} g_i f_i = 1 \qquad ①$$

由 Brownawell(1987) 和其他人建立的有效零点命题说到在上述命题中,人们可以假定 g_i 的次数满足

$$\deg g_i \leqslant \max(3,D)^n, D = \max \deg f_i$$

以这个次数为囿界,这个判定问题可以变成一个线性代数的命题.给定 f_i 的系数,人们可以检验 ① 对于 g_i 的系数是否有解,因此人们可以有一个判定零点定理的算法.如此算法所需的步数是以(作为输入的大小的)f_i 系数数目指数增长的.

(在C上的)猜测　不存在判定复数域 $\mathbf{C}$ 上的 Hilbert 零点定理的多项式时间算法.

一个多项式时间算法是这样一种算法,它的步数由诸 f_i 的系数数目的一个多项式所界定.

为了数学地理解这个猜测,人们需要算法的严格定义.在此范围内,传统的 Turing 机的定义是无意义的.

非常简要地说,$\mathbf{C}$ 上的机器是以复数的有穷串作为输入、状态和输出的,另外,关于状态的计算包括算术运算和串移送.最后还要有关于"$x_1 = 0$ 成立否"的分支运算.

输入的大小是指输入串中元素的个数.计算时间指在由输入到输出的过程中所用的机器运算的次数.因此 $\mathbf{C}$ 上多项式时间算法是意义明确的.

应该指出关于机器的猜测的陈述中只用到 $\mathbf{C}$ 的结构是一个域,因此这里的机器和猜测对任何域都有效.特别地,当域是二元素的 $\mathbf{Z}_2$ 时,我们得到 Turing 机的定义.

考虑判定问题:以系数在 $\mathbf{Z}_2$ 中的 n 个变元的 k 个多项式为输入,是否存在一公共零点 $\zeta\in(\mathbf{Z}_2)^n$?

猜测:不存在可判定此问题的 $\mathbf{Z}_2$ 上多项式时间算法.

这就是经典猜测 $\mathrm{P}\neq\mathrm{NP}$ 的简单重述.

在上面我们回避了有关 NP 完全性的基本概念和定理.Cook 和 Karp 的经典情况可参见[Garey-Johnson,1979],任意域上理论的经典情况可参见 BCSS.

问题 4　多项式的整数零点

我们首先定义一个源于复杂性理论的丢番图不变量 τ.对于一个整系数单变量的多项式 $f\in\mathbf{Z}[t]$,它的一个程序是如下定义的"对象"$(1,t,u_1,\cdots,u_k)$:

$u_0 = t, u_{-1} = 1, u_k = f$,使得对任何 $l, u_l = u_i \circ u_j (i, j < l)$,这里 $\circ$ 表示 $+$, $-$, $\times$ 中的任一种运算.定义 $\tau(f)$ 为在所有这样的程序中 k 的最小值.

问题:f 不同整零点的个数是否能被 $\tau(f)$ 的一个多项式所界定?换言之,是否对所有的 $f \in \mathbf{Z}[t]$

$$Z_a(f) \leqslant \tau(f)^c$$

这里 $Z_a(f)$ 表示 f 不同整零点的个数,c 是一个与 f 无关的常数.

我和 Mike Shub 在研究复杂性问题时发现了上述问题.我们证明了该问题的肯定回答,将意味着零点理论作为 $\mathbf{C}$ 上的一个决策问题是难以驾驭的,因此在 $\mathbf{C}$ 上 $\mathrm{P} \neq \mathrm{NP}$.

由于 f 的次数不超过 $2^{\tau(f)+1}$,它的全部零点数将不会超过 $2^{\tau(f)+1}$.

对于 Chebyshev 多项式,其不同实零点的个数就呈 τ 的指数级增长.

许多经典的丢番图问题是关于两个或更多个变量的.这里所问的仅仅是单个变量时的一个估计,然而,它似乎已经不太容易.

此外,还有一个与之相关的问题.对一个整数 m 可类似定义它的程序$(1, m_1, \cdots, m_l)$:令 $m_l = m, m_0 = 1$ 使得 $m_q = m_i \circ m_j, i, j < q$,这里 $\circ$ 表示 $+$, $-$, $\times$ 中任一种运算.在 m 的所有程序中,$\tau(m)$ 表示 l 的最小值.因此,$\tau(m)$ 代表了从 1 开始,通过加、减、乘得到 m 的最短过程.

问题:是否存在一个常数 c,使得对所有整数 k,$\tau(k!) \leqslant (\log k)^c$,人们也许期望这是错的,以致 $k!$ 是"难计算的".

问题 5:丢番图曲线高度的界

是否存在一个与 f 无关的常数 c,使得丢番图方程 $f(x, y) = 0$(输入 $f \in Z[u, v]$) 的可行性能在时间 2^{s^c} 内被确定?

这里 $s = s(f)$ 是 f 的"大小",其定义如下

$$s(f) = \sum_{|\alpha| \leqslant d} (\log |a_\alpha| + 1), f(x, y) = \sum_{\alpha \leqslant d} a_\alpha x^{\alpha_1} y^{\alpha_2}$$

$$\alpha = (a_1, a_2) \text{ 及 } |\alpha| = \alpha_1 + \alpha_2, \alpha_i \geqslant 0$$

此外,如果存在整数 x, y 使得 $f(x, y) = 0$,则称 f 是可行的.我们假定这里使用的是 Turing 计算模型.

上述问题基本上是在[Cucker-Koiran-Smale, 1997]提出的.$s(f)$ 实际上是 f 的某种形式的"高度".例如,[Lang, 1994]中所猜测的高度的界对解决此问题可能会有帮助.也可参见[Manders-Adleman, 1978].

问题 6:天体力学中相对平衡态数目的有限性

给定正实数 $m_1, \cdots, m_n$ 作为天体力学的 n 体问题中物体的质量,问相对平衡态的数目是否有限?

此问题见于 Wintner 关于天体力学的书(1941)中.一个相对平衡态是指由





平面旋转诱导的牛顿方程的一个解.

三体问题有五个相对平衡态:三个由 Lagrange 发现,另两个由 Euler 发现.四体问题(相对平衡态)的有限性尚不清楚.

在文[Smale,1970]中,我说明了相对平衡态是平面 n 体问题中由势能诱导的函数的临界点,确切地说,相对平衡态对应于

$$\hat{V}:(S_k-\triangle)/SO(2)\to\mathbf{R}$$

的临界点,其中 $S_k=\{x\in(\mathbf{R}^2)^n\mid\sum m_ix_i=0,\frac{1}{2}\sum m_i\|x_i\|^2=1\}$,$\triangle=\{x\in S_k,x_i=x_j$ 对于某个 $i\neq j\}$.旋转群 $SO(2)$ 作用在 $S_k-\triangle$ 上,$\hat{V}$ 是由势能函数

$$V(x)=\sum_{i<j}\frac{m_im_j}{\|x_i-x_j\|}$$

在商空间诱导得到的.注意到 $V:S_k\to\mathbf{R}$ 在旋转群 $SO(2)$ 的作用下是不变的,并且商空间 $S_k/SO(2)$ 同胚于 $n-2$ 维复射影空间.

Mike Shub(1970)证明临界点的集合是紧的,Palmore(1976)证明即使 $n=4$ 时,$\hat{V}$ 也可能有退化的临界点.

相对平衡态在天体力学中起着重要的作用,例如,它在角动量映射产生的分支中的作用.此外在太阳系中存在的"*Trojans*"("脱罗央群",是一群小行星——译注),就相应于 Lagrange 相对平衡态.

Kuz'mina(1977)对于通有情形找到了明确的上界.

进一步的背景材料可以在[Abraham-Marsden,1978]中找到.

问题 7:2 维球面上点的分布

令 $V_N(x)=\sum_{1\leq i\leq j\leq N}\log\frac{1}{\|x_i-x_j\|}$,其中 $x=(x_1,\cdots,x_N)$,x_i 是 2 维球面 $S^2\subset\mathbf{R}^3$ 上的不同点,$\|x_i-x_j\|$ 是 $\mathbf{R}^3$ 中的距离.记 $\min_xV_N(x)$ 为 V_N.

寻找$(x_1,\cdots,x_N)$,使得

$$V_N(x)-V_N\leq c\log N,\text{其中 } c \text{ 为一通用常数} \qquad ②$$

"寻找"是指给出一个算法,输入 N,就会输出 2 维球面上的满足②的 N 个不同的点 $x_1,\cdots,x_N$.准确地,我们可以借助关于 N 的停机时间多项式得到 BCSS 意义上的实数算法(添加一个平方根计算).

这个问题来自于我跟 Mike Shub 一起对复杂性理论的研究[Shub-Smale,1993],动机是寻找一个关于同伦算法的好的起始多项式,以深入了解现代数学基本定理.

满足 $V_N(x)=V_N$ 的点$(x_1,\cdots,x_N)=x$ 称为 N 元椭圆 Fekete 点(见[Tsuji,1959]).

作为 N 的函数,V_N 满足

$$V_N = -\frac{1}{4}\log\left(\frac{4}{e}\right)N^2 - \frac{1}{4}N\log N + O(N)$$

很自然地也可以考虑函数

$$V_N(x,s) = \sum_{i<j}\frac{1}{\|x_i - x_j\|^s}, V_N = \min V_N(x,s)$$

x 同上,$0 < s < 2$.当 $s = 0$ 时,自然地对应着原来的 $V_N(x)$,V_N.当 $s = 1$ 时,$V_N(x,1)$ 是 Coulomb 势函数,$V_N(1)$ 对应着被约束在 2 维球面上的 N 个电子的某个平衡位置.

对以上主要问题我曾求助于 Ed Saff.后来,他和他的同事们在这个主题和相关邻域写出了许多好文章.这方面的背景和进一步的资料可参看[Kuijlaars-saaf,1997}和[Saff-Kuijlaars,1997].而在[Rakhmanov-Saff-Zhou,1994}中,可以找到支持这些问题的数值证据($N = 12\,000$).

求解我们的主要问题的另一个办法是优化函数

$$W_N(x) = (\exp V_N(x))^{-1} = \prod_{i<j}\|x_i - x_j\|$$

然而正像在[Shub-Smale,1993]中所写到的,"…… 这个办法不会是容易的,因为有指标 N 的鞍点(在 S^2 的大圆上平均散布着 N 个点 $x_1,\cdots,x_N$),而且 W_N 具有的各种对称性也会搅乱图象."

问题 8:把动力学引进经济理论中

下面的问题不是纯数学问题,它位于经济学和数学交汇处.目前,人们只是对这问题的极特殊情形给出了解决.

把一般的均衡理论的数学模型推广到包含价格调节的情形.

经济学中有一个关于平衡价格的(静态)理论,它为 Walras 所开创,由 Arrow 和 Debreu 奠定了坚实的基础(见[Debreu,1959]).在只有一个市场的平凡情形,这相当于一种均衡:"供给等于需求",而且很容易获得一种常态的动力学[Samuelson,1971].对于多个市场,情况则非常复杂.

有一个从价格空间到商品空间的过剩需求函数 $Z(p) = D(p) - S(p)$.需求 D 和供给 S 都定义为个体的总和.经济学理论认为涉及个体行为的条件是正当的,它们引导出 Z 上的公理.对于过剩需求映射 $Z:\mathbf{R}_+^l \to \mathbf{R}^l$ 的这些公理为:

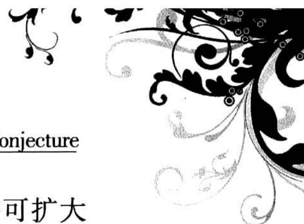
(1) $Z(\lambda_p) = Z(p), \forall p = (p_1,\cdots,p_l), p_i \geqslant 0, \lambda \in \mathbf{R}, \lambda > 0$.

(2) $\sum_{i=1}^{l}, p_i Z_i(p) = 0$,Walras 法则(总值为零).

(3) 如 $p_i = 0$,则 $Z_i(p) > 0$(对免税商品为正需求).

由(1),(2),(3),Z 可解释为$(l-1)$维球面与正象限(positive orthant)的交集上的向量场,该向量场在边界上指向内部.均衡价格向量 p^* 的存在性可由 Hopf 定理得出,所以 $Z(p^*) = 0$,及"供给等于需求".





问题 8 要求一个动力学模型来描述,其状态变量为价格向量(或许可扩大到包含其他经济变量).这个理论应该与现有的均衡理论相匹配.最希望得到的特征是由经济的个体行为所决定的价格随时间的演化.

我在这个问题上已工作了几年,感觉到它是经济理论中的主要问题[Smale,1976],了解其背景可参见[Smale,1981].

问题 9:线性规划问题

是否存在基于实数的多项式时间算法来决定不等式线性系统 $A_x \geqslant b$ 的可行性?

这个问题所要求的算法是由一个实数机器给出的 BCSS 意义上的算法(见问题 3).系统 $\boldsymbol{Ax} > \boldsymbol{b}$ 输入一个 $m \times n$ 实矩阵 $\boldsymbol{A}$ 和一个向量 $\boldsymbol{b} \in \mathbf{R}^m$,问题为:是否存在 $\boldsymbol{x} \in \mathbf{R}^n$,使得对所有 $i = 1, \cdots, m$,都有 $\sum_{j=1}^{n} = a_{ij}x_j \geqslant b_i$?时间由算术运算的次数来决定.这个问题出现在 BCSS 中.

这是线性规划中优化问题的最重要的一种提法.如上给定 $\boldsymbol{A}$, $\boldsymbol{b}$ 和 $\boldsymbol{c} \in \mathbf{R}^n$,判定满足条件 $\boldsymbol{Ax} \geqslant \boldsymbol{b}$ 的

$$\max_{x \in \mathbf{R}^n} \boldsymbol{c} \cdot \boldsymbol{x}$$

是否存在,如果存在,输出一个这样的 $\boldsymbol{x}$.

Dantzig 著名的单纯形法对于($\mathbf{R}$ 上的)这两个问题提供了(基于实数的)算法.但是 Klee 和 Minty 证明了这个算法在最坏的情形其运算速度是指数式的缓慢.另外,Borgwardt 和我在 Haimovich 的极为重要的帮助下证明了大多数情况下它是多项式时间的.关于这些可参看[Schrijver,1986].

根据计算的 Turing 模型,利用有理数域 $\mathbf{Q}$,而且由"比特"(bits)来衡量的话,则有平行的理论.起始于 Rudin-Nemirovsky 的思想,Khachian 对线性规划问题找到了一个多项式时间算法(椭球法).后来 Karmarkar 借助于他的"内点法"找到了一个实际的算法,可以在 Turing 模型上以多项式时间运行.对于这些,可参看文献[Grotschel-Lovász-Schrijver,1993]和[Schrijver,1986].

与基于 $\mathbf{R}$ 上的上面的主要问题密切相关的一个类似的问题是:对这些线性规划问题是否存在基于 $\mathbf{Q}$ 上的"强多项式算法"?这个算法要求算术运算的次数和节运算(Bitoperation)的次数都是输入次数($m \times (n+1)$)的多项式.Megiddo,特别是 Tardos 对这个问题已给出部分解答.(见[Grotschel-Lovász-Schrijver,1993]).

而对于 $\mathbf{R}$ 上的问题,也有参考文献[Barvinok-Vershik,1993]和[Traub-Wozniakowski,1982].

问题 10:封闭引理

设 p 是紧致流形 M 上的一个微分同胚 $S: M \to M$ 的非游荡点.S 是否可以

直到 r 阶导数(对于每一个 r) 被 M 上的微分同胚 $T:M \to M$ 任意好地逼近 ($C^{r'}$ – 逼近),使得 p 是 T 的周期点?

非游荡点 $p \in M$ 是指 P 具有如下性质:对于 p 的每个邻域 U,存在 $k \in \mathbf{Z}$ 使得 $S^kU \cap U \neq \varnothing$. 这里是 S^k 是 S 的 k 次迭代;如果 $T^m(p) = p$, p 是 T 的周期为 m 的周期点.

这就是“封闭引理”的离散形式,它的 C^1 情形已由 Charles Pugh(1967) 给出肯定的解答.

很容易找到满足要求的 C^0 – 逼近. Peixoto 发现这一论证对于 C^1 不成立,修正了 René Thom 的一个错误(René 告诉我这是他犯的最大的错误).

Pugh 和 Rohinson(1983) 就 Hamilton 情形证明了 C^1 – 逼近的封闭引理. Peixoto 正面给出了圆周上以及可定向二维流形上连续时间情形的 C^r – 逼近(对任意 r). 最近通过 Hayashi(1997) 的工作封闭引理被赋予了进一步的重要性;也见[Wen-hia, 1997].

问题 11:一维动力学是否通常是双曲的?

一个复多项式 T 能否被具有这样性质的同次多项式所逼近,其每一个临界点在迭代下趋向于某个周期汇(periodic sink)?

即使对于二次多项式这一问题也还没有解决. 这里考虑由多项式映射 $T:\mathbf{C} \to \mathbf{C}$ 的迭代生成的离散动力系统. 因此若 $z \in \mathbf{C}$,它的时间轨道 $z = z_0, z_1, z_2, \cdots$ 由 $z_i = T(z_{i-1})$ 定义,i 可理解为(离散的) 时间. T 的不动点 $w(T(w) = w)$ 是一个汇,如果 T 在 w 处的导数 $T'(w)$ 的绝对值小于 1, T 的周期为 p 的周期汇是 T^p 的一个汇. T 的临界点是指 T 的导数为 0 的点.

精确地阐述这一问题后,从产生于 20 世纪 60 年代的双曲动力学的框架去看是很有帮助的. 微分同胚 $T:M \to M$ 的不动点 x 是双曲的,如果 T 在 x 处的导数 $DT(x)$(作为切空间的线性自同构) 没有绝对值为 1 的特征值. 若 x 是周期为 p 的周期点并且是 T^p 的双曲不动点,那么 x 是双曲的. 双曲性的定义可以自然地扩展到 Ω—— 非游荡点集的闭包上(见问题 10).

一个动力系统 $T \in \mathrm{Diff}(M)$ 称作是双曲的(或满足公理 A),如果周期点集在 Ω 中稠密,并且 Ω 是双曲的[Smale, 1967] 或[Smale, 1980]. 我们还要假定一个无环条件. 许多人特别是 Ricardo Mané 的工作,将双曲动力学与被称做结构稳定性的动力学稳定性的一个很强的概念等同起来. 这甚至是这一类动力系统结构理论研究的开始.

双曲系统构成了动力系统的一个大集合,但是包含应用混沌动力系统在内的一个更大的集合却超出了它的范围. 双曲性的概念从可逆动力系统扩展到我们上述问题的情形,即从 $\mathbf{C}$ 到 $\mathbf{C}$ 的多项式映射. 经典的复变理论允许我们重新提出一个等价的问题:





多项式映射 $T:\mathbf{C}\to\mathbf{C}$ 能否被双曲多项式映射所逼近?

复一维动力系统理论在20世纪初为Fatou和Julia所创立.在20世纪60年代我让我的博士生John Guckenheimer看这一领域的文献并试图解决上述问题(和其他一些问题).他的论文(见[Chern and Smale,1970])包含了一个肯定的回答,但在证明中有一个漏洞.现在这一问题作为一维动力系统的基本问题仍未解决.

John的带有错误证明的文章是自Poincaré以来许多关于这一问题的许多文章中的一篇.

复一维动力学成为一个活跃的研究邻域,其中包含了Douady,Hubbard,Sullivan,Yoccoz,McMullen以及其他人的重要贡献.(见[Mcmullen,1994])

对于由光滑映射 $T:I\to I, I=[0,1]$ 定义的实一维动力学有着平行的研究邻域.

问题:光滑映射 $T:[0,1],[0,1]$ 能否被双曲映射 C^r - 逼近(对所有 $r>1$)?

在Guckenheimer的论文的同一时期,我让Ziggy Nitecki研究这一问题.我最初的忽视导致了没能发现Nitecki论文的错误(见[Chern and Smale,1970]),他试图给出上面问题的一个肯定的答复.

其后,Jakobson(1971)就 C^1 - 逼近回答了上述问题,但一般情形仍未解决,背景材料见[de Melo-van Strien,1993].

问题12:微分同胚的中心化子

紧流形 M 到它自身的微分同胚是否能由只与其迭代可交换的 $T:M\to M$ 所 C^r 逼近(对所有的 $r\geqslant 1$)?

于是 T 在微分同胚群 $\mathrm{Diff}(M)$ 中的中心化子应该为 $\{T^k \mid k\in\mathbf{Z}\}$.

在[Smale,1963]中我就开始考虑中心化子的问题,但直到Nancy Kopell跟随我写出学位论文之后(见[Chern-Smale,1970]),在 $\dim M=1$ 的情形问题才得到肯定的回答.这个问题是我提出来的[Smale,1967],今天对于2 - 维流形 M,问题仍然没有解决.

人们可能会问,具有平凡中心化子的 M 的所有微分同胚之集在 C^r 拓扑之下在 $\mathrm{Diff}(M)$ 中是否是稠密和开的?

对这些问题的主要贡献是由Palis-Roccoz(1989)做的,对任何流形的双曲动力学情形,问题几乎得到了完全解决(见问题11).

在[Smale,1991]中我写到:“我发现这个问题的有趣之处在于它对于双曲以外的完全未知的领域注入了曙光,在那些邻域要想清晰地提出这个问题都是很难的.”

问题13:Hilbert第十六问题

考虑 $\mathbf{R}^2$ 上的微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t}=P(x,y),\frac{\mathrm{d}y}{\mathrm{d}t}=Q(x,y) \qquad ③$$

其中 P 和 Q 是多项式.系统的极限环的数目是否存在上界 K 具有形式 $K\leqslant d^q$, d 是 P 和 Q 的阶数的最大值,q 是一个普适常数?

这是 Hilbert 第十六问题第二部分的新近的提法.除了 Riemann 假设,它看起来是 Hilbert 问题中最难以捉摸的一个.

事实上,自从 Petrovskiĭ 和 Landis(1957) 的文章企图给出一个正面的回答之后,这一问题的进展似乎是倒退的,早期,Dulac(1923) 宣称系统(3) 具有有限数目的极限环.在 Petrovskiĭ 和 Landis 的文章被发现了一个漏洞后(见[Petrovskiĭ and Landis,1959]),Ilyashenko(1985) 发现了 Dulac 文章中的一处错误.进一步史松龄(1982) 找到了 Petrovskii 和 Landis 关于 $n=2$ 情形给出的特定的上界的一个反例.接着,Ecalle 和 Ilyashenko 各自独立地给出 Dulac 断言的证明的两篇文章面世了,这两篇文章[Ecalle,1992] 和[Ilyashenko,1991] 尚未完全被数学界所消化.

这样有限性得到了,但还没有上界.我们将考虑一类特殊的方程,它的有限性是简单的,但有界性尚未证明.

下述方程相应于 Liénard 方程(可见[Hirsch and Smale,1974])

$$\frac{\mathrm{d}x}{\mathrm{d}t}=y-f(x),\frac{\mathrm{d}y}{\mathrm{d}t}=-x \qquad ④$$

其中 $f(x)$ 是首项为 x^{2k+1} 的实多项式且满足 $f(0)=0$.

若 $f(x)=x^2-x$,那么④是具有一个极限环的 van der Pol 方程.更一般地,很容易证明 ④ 的所有解顺时针方向环绕平衡点(0,0).沿着这些解可以定义"Poincaré 截面"$T:\mathbf{R}^+\to\mathbf{R}^+$,其中 $\mathbf{R}^+$ 为正 y 轴.④ 的极限环恰好就是 T 的不动点.我在各个不同的演讲中都提及了估计这些不动点数目的问题.作为回应,Linz,de Melo 和 Pugh(1997) 发现了具有 k 个不同极限环的例子,并猜测 k 就是极限环数目的上界.$(\deg f)^q$ 形式的上界仍未被发现.由于 T 是解析的,对每个 f,④ 有有限个极限环.

进一步的背景材料见[Browder,1976],[Ilyashenlo and Yakovenko,1995],[Lloyd and Lynch,1988] 和[Smale,1991].

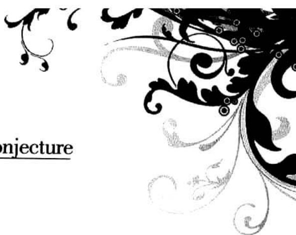
问题 14:Lorenz 吸引子

Lorenz 常微分方程的动力学是否就是 Williams,Guckenheimer 和 Yorke 的几何 Lorenz 吸引子的动力学?

Lorenz 方程是

$$\dot{x}=-10x+10y$$
$$\dot{y}=28x-y-xz$$





$$\dot{z} = -\frac{8}{3}z + xy$$

Lorenz(1963) 利用计算机分析上述方程时发现它的大多数解趋于一个确定的吸引集,从而他得到了一个重要的"混沌"的早期实例,然而缺少数学证明.这一数值工作刺激了看起来具有同样行为的由几何定义的常微分方程的严格的数学发展(Yorke,[Williams,1979],[Guckenheimer and Williams,1979]).几何吸引子已被分析得很透彻,差不多可以说它已被完全理解了.

问题 14 是问最初方程的动力学是否与几何模型的动力学一致.最完全的解答将是描述一个 $\mathbf{R}^3$ 到 $\mathbf{R}^3$ 的同胚,它将 Lorenz 方程的解映到几何吸引子——或者说,几何吸引子双参数族中某一吸引子的解(解族中的一些解).

这一问题的解答将是建立应用动力系统理论基础的第一步.时至今日,在工程和物理的方程中,混沌还只是在某种弱的意义下被建立起来,即证明马蹄的存在性(如 Melnikov,Marsden,and Holmes; 见[Guckenheimer and Holmes, 1990]).

动力系统中几何的、结构稳定的混沌吸引子出现在我的文章中[Smale, 1967].但这些吸引子并没有在任何物理系统中产生.

问题 14 的一些部分结果属于 Richlik 和 Robinson,见[Robinson.1989].

问题 15:Navier-Stokes 方程

$\mathbf{R}^3$ 中 3 维域 Ω 上的 Navier-Stokes 方程有唯一的整体光滑解吗?

这也许是偏微分方程中最著名的问题.说得更确切些,Navier-Stokes 方程可以写成

$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u - v\Delta u + \text{grad } p = 0, \text{div } u = 0$$

要找满足这些方程的 C^∞ 映射 $u:\mathbf{R}^+ \times \Omega \to \mathbf{R}^3, p:\Omega \to \mathbf{R}$,在 $t = 0$ 及边界 $\partial\Omega$ 上 u 值给定.这里 $\mathbf{R}^+ = [0, +\infty)$,$u \cdot \nabla$ 是算子 $\sum_1^3 u_i \frac{\partial}{\partial x_i}$,$v$ 是正常数.细节可参看[Chorin Marsden,1993].

许多数学家研究过这个问题.二维与三维情况当 t 在小时间区间$[0, T]$时已有肯定答案.参看[Temam,1979].

很有理由认为这个问题的解是搞清楚湍流这个大问题的根本的一步.例如这可以帮助领悟[Ruelle and Takens,1971]的思想,这个思想把混沌吸引子的概念引入湍流模型中去.也参看[Chorin,Marsden and Smale,1977].

在[Smale,1991]中,我曾问过环面上带强迫项的 2 维 Navier-Stokes 方程的解当时间趋于无穷时是否一定收敛到平衡态.Babik 和 Vishik(1983) 已经给出一些反据.后来 Liu(1992) 给出证明了收敛到更复杂的吸引子的一些例子.

问题 16:Jacobi 猜想

假定 $f:\mathbf{C}^n \to \mathbf{C}^n$ 是多项式映射,它有如下性质:在每点的微分都是非奇异的,那么 f 是否一定是一一对应的?

这里 $f(z) = (f_1(z),\cdots,f_n(z))$,每个 f_i 是一个 n 变元多项式.f 在 z 的微分 $Df(z):\mathbf{C}^n \to \mathbf{C}^n$ 可以考虑为偏微分做成的矩阵,非奇异条件是 $\det Df(z) \neq 0$.

若 f 确是内射的,那么 f 一定是满的,并有逆,它是一多项式映射,对此有一初等证明,请看[Rudin,1995].

此问题可回溯到 20 世纪 30 年代.请参看[Bass,Connell 和 Wright,1982]和[van der Essen,1997]关于此问题的重要性、背景及相关结果的精彩的综述.

问题 17:解多项式方程组

对于 n 个 n 变元复多项式方程,能否有一种统一算法,平均用多项式时间近似地找到一个零点?

在我与 Mike Shub 合作的五篇系列文章[Shub 和 Smale,1994]中的最后一个定理即与上述结果除了"统一"一词之外全相同.

我们复习一下定义.考虑 $f:\mathbf{C}^n \to \mathbf{C}^n, f(z) = (f_1(z),\cdots,f_n(z)), z = (z_1,\cdots,z_n)$,其中 f_i 是 d_i 次 n 变元多项式.通过加上一个变元 z_0 可以合理地将 f_i 变成齐次多项式,从而可在相应的射影空间中工作,然后再将算法和结果还原到原来的仿射空间中.

"近似地"本质上可定义为使用牛顿方法.根据 Abel 和 Galois 及其他人的结果知道近似是必要的.如果你想要得到关于时间的形式的定义,可以说它是由在真正的机器上所作的算术运算和比较"$\geqslant$"的次数来度量的(如问题 3 中所说的).

概率的测量必须对每个 $d = (d_1,\cdots,d_n)$ 赋予所有这样的 f 的函数空间上,算法的时间在 f 的这个空间上被平均.是否存在这样一个算法,使得它的平均时间被 f 的系数个数(输入大小)的多项式所界定?

在[Shub and Smale,1994]中证明了这可以做到,但对不同的 d,甚至对于每个所希望的成功概率的算法都是不同的.统一算法是与 d 无关的(d 是输入的一部分)的一种算法.

自然,找到多项式或多项式组的零点是数学中最古老和最中心的问题之一.本猜想是问,是否在问题中所指定的条件下可用计算机系统地解开.如果不存在多项式时间方式求零点,则任一计算机都无法获得成功.

将多项式零点问题放在一般情况下考虑近来颇有点进展.Hilbert 零点定理(作为一个判定问题)在任何域上都是 NP - 完全的(见问题 3).

此问题也可对于实数域提出,问题是类似的,但更困难.

问题 18:智能的极限

人工的和人类的智能,两者的极限是什么?





附录Ⅵ　Poincaré 猜想和三维流形分类的近期进展[①]

引　言

受当时关于曲面的熟知事实的启发,在对三维流形作了大量拓扑研究之后,Poincaré 在他的 1904 年论文的结尾处陈述还剩下一个问题要处理,即(用现代语言重述):

如果 M 是一个基本群平凡的闭三维流形,那么 M 微分同胚[②]于 S^3 吗?

Poincaré 猜想是说,这个问题的答案是"对".发展工具以攻克这个问题形成了 20 世纪三维拓扑的许多工作的基础,包括例如 Dehn 引理和圈定理(the loop theorem)的证明,以及纽结和链环上手术的研究.

在 20 个世纪 80 年代,Thurston 发展了对三维流形的另一种处理方法,他考虑带常负曲率 -1 的 Riemann 度量的三维流形,这些局部等距于三维双曲空间的流形称为双曲流形.存在相当明显的阻碍显示并非每个三维流形都能有这样一个度量.Thurston 提出了一个一般的猜想,粗略地说就是,这些明显的阻碍是仅有的阻碍;如果对一个特殊的三维流形这些阻碍消失,则那个流形容许这样的度量.他对这个猜想的各种重要特殊情形的证明导致他提出了一个关于所有(三维——原译注)流形上双曲或其他局部齐性度量的存在性的更一般的猜想;这称为 Thurston 的三维流形几何化猜想.这个猜想的陈述有些复杂,因此将把它往后推到第 2 节.一个重要之处是 Thurston 的几何化猜想包含 Poincaré 猜想作为一个非常特殊的情形.另外,Thurston 的猜想在两点上优于 Poincaré 猜想:

- 它对所有闭定向三维流形都适用.
- 它断定在三维时拓扑和几何之间有密切关系.

事实上,在我看来,在 Thurston 关于三维双曲流形的工作和他对一般的几何化猜想的阐述之前,专家对 Poincaré 猜想是对还是错意见并不一致.在 Thurston 的工作之后,尽管它对 Poincaré 猜想并无直接影响,人们一致认为 Poincaré 猜想(和几何化猜想)应该是对的.把 Poincaré 猜想纳入一个更广的猜

---

① 原题:Recent Progress on the Poincaré Conjecture and the Classification of 3-manifolds.
译自:Bull.(New Series) of the AMS, Vol.42(2005),No.1,p.57-78.

② 每个拓扑三维流形都有一个微分结构,且它在不计接近于恒同映射的微分同胚的意义下是唯一的.因此,对三维流形分类相差同胚和相差微分同胚是等价的.——原注

想，然后给这个更广的猜想提供证据（独立于 Poincaré 猜想），导致了对 Poincaré 猜想的更坚定的信念，这有点似非而是. 关于 Poincaré 猜想的历史、三维流形拓扑的发展和 Thurston 几何化猜想的更多细节参看 Milnor 的综述文章.

Thurston 几何化猜想与我们最相关的方面是，三维流形上特别好的度量的（猜想的）存在暗示对三维流形分类问题的一种更分析化的处理方法. 通过在（一个任意维流形上的）Riemann 度量的空间上引入瑞奇（Reccl）流①，Hamilton 提出了这样的一种处理方法. 然后他猜想从一个紧三维流形上的任意 Riemann 度量开始，瑞奇流将产生一个单参数族的度量，趋于 Thurston 几何化猜想所假设的好的度量. 但是这个纲领存在很多技术问题：我们知道一般说来瑞奇流将在有限时间内产生奇点，因此必须找到分析这些奇点和延拓瑞奇流越过奇点的方法. 进一步，即使瑞奇流延拓到了所有时间，仍然剩下关于在时刻 $t = +\infty$ 极限的性质的复杂问题.

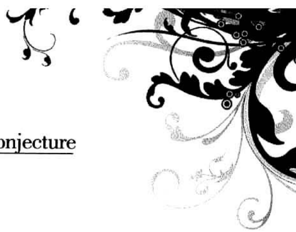
为了理解度量和曲率的演化，Hamilton 建立了许多关键的分析估计，由此对他首先提出的纲领作出了有意义的进展，他还通过应用 Cheeger-Gromov 的几何极限论据证明，在特殊情形人们可以利用瑞奇流来建立 Thurston 的几何化猜想. 更精确地说，在某些条件（包括瑞奇流在有限时间内不产生奇点）下，他能够证明，当 $t \to \infty$ 时（所考虑的）这族度量经过尺度变换可以收敛到一个完备双曲度量；因此所考虑的流形有一个双曲度量. 最近，在 Hamilton 工作的基础上，Perelman 在一系列（3 篇）预印本中断言他已经克服了 Hamilton 纲领的所有各种技术的、几何的和分析的困难. Perelman 的结论的基础从对瑞奇流中大曲率区域的详细分析开始，他对这些区域给了一个好的分类，这足以使他分析瑞奇流中产生的有限时间奇点并给出它们的粗略分类. 利用这个他展示如何通过自然的方式扩张瑞奇流越过奇点，构造他所说的带手术的瑞奇流.（Hamilton 以前在四维流形中定义了一种类似的手术过程）Perelman 已经把 Hamilton 的许多关于瑞奇流的估计扩充到带手术的瑞奇流. 通过利用这些技术. Perelman 能够对所有正的时间延续瑞奇流和手术过程. 最后，他把 Hamilton 关于 $t \to \infty$ 时极限的分析扩充到带手术的瑞奇流. 这样，他断言他已经建立了 Thurston 几何化猜想，从而也证明了 Poincaré 猜想. Perelamn 的论证非常复杂，涉及许多惊人的原创性思想，数学界仍在试图消化他的论证，查明它是否真正完备、正确. 截至 2004 年 8 月，这些论证经受住了检查. ②尽管对他的论证的分析尚未完成，显然 Perelman 已经对理解三维流形上瑞奇流的性质做出了巨大的进展.

---

① 关于度量的瑞奇流在第 3 节定义.——原注

② 关于更新的情况可参见曹怀东 – 朱熹平在 Asian J. Math. 10(2006), No. 2 上的论文和 Kleiner-Lott 及 Morgan – 田刚在 arXiv 上的预印本.——原译注





1 局部齐性流形

让我们搭起 Hamilton 的 Perelman 的工作舞台.

1.1 齐性几何和几何结构

如果 Riemann 流形(M,g)的等距群传递地作用在它上面,就称它为齐性 Riemann 流形.这意味着这个流形看起来在任意点度量都一样.齐性流形的例子有圆球面 S^n,欧氏空间 $\mathbf{R}^n$,和双曲空间 H^n.后者是 $\mathbf{R}^{n+1}$ 中由$\{Q(x)=-1\}$给出的超曲面(的上叶 —— 原译注),这里 $Q(x_0,x_1,\cdots,x_n)=-x_0^2+x_1^2+\cdots+x_n^2$,其 Riemann 度量由 Q 的其切平面上的限制给出.Q 的正交群的指数为 2 的子群(由保持双曲面的上叶不变的变换组成)作为 H^n 的全等距群传递地作用在 H^n 上.H^n 的所有截面曲率都等于 -1.

我们说一个 Riemann 流形以一个给定的齐性流形(M,g)为模型,如果这个流形的每个点都有一个邻域等距于(M,g)的一个开集.这样的流形如果完备的话就称为局部齐性流形,一个局部齐性流形的万有覆盖(带着其诱导 Riemann 度量)是它以之为模型的齐性流形.我们主要关心体积有限的局部齐性流形.

在二维,在等距意义下有4个单连通齐性流形:S^2,$\mathbf{R}^2$,H^2 和 G(带有左不变度量),这里 G 是群 $\mathbf{R}\times\mathbf{R}^*$,其中 $\mathbf{R}^*$ 自然作用在 $\mathbf{R}$ 上.然而没有面积(原文为 volume—— 译注)有限的局部齐性流形以第 4 个例子为模型,因此我们只关心前 3 个模型.第 1 类有限面积局部齐性流形是球面和射影平面,第 2 类是环面和 Klein 瓶,第 3 类是紧或非紧的有限面积完备双曲曲面.当然,对后两类来说,其上存在不等距的局部齐性度量的连续族.这给我们以下熟知和经典的定理:

定理 1.1.1(二维单值化) 设 X 是紧曲面,则 X 容许一个局部齐性度量,局部等距于上面的常曲率模型中的一个.如果 $\chi(X)>0$,则模型有正曲率;如 $\chi(X)=0$,则平坦;如 $\chi(X)<0$,则有负曲率(即为双曲的).

1.2 三维

在三维,每个有限体积的局部齐性流形都以下面列出的 8 种齐性流形为模型(如二维时一样,存在其他齐性三维流形,但不存在有限体积局部齐性流形以它们为模型).首先,我们有常(截面)曲率的例子:

(1)S^3,常曲率 $+1$.

(2)$\mathbf{R}^3$,平坦.

(3)H^3,常曲率 -1.

任何以 S^3 为模型的局部齐性流形都有常正截面曲率.它们形如 S^3/Γ(这里 Γ 是$SO(4)$ 的有限子群,自由地作用在 S^3 上),包括 S^3,RP^3,透镜空间,以及 S^3 在正多面体的对称群作用下的商.这些流形称为球面空间型.注意所有球面空间型都是紧的.

紧平坦三维流形的原型是平坦环面.这些是 $\mathbf{R}^3$ 在由平移构成的完全(即秩 3)格作用下的商.任何以 $\mathbf{R}^3$ 为模型的局部齐性流形都有平坦度量.如果它的体积有限,则它被平坦 T^3 有限覆盖,从而自动紧.

以 H^3 为模型的局部齐性流形是完备双曲流形,即所有截面曲率都为 -1 的完备 Riemann 流形.存在无穷多个互不微分同胚的紧或非紧的有限体积的这种流形.由 Mostow 刚性定理,一个三维流形最多容许一个(在等距意义下)完备、有限体积双曲度量.如果考虑的流形非紧,则无穷远点的邻域的每个分支都微分同胚于 $T^2\times(1,\infty)$,这些环面切片几何上随着 $t(\in(1,\infty)$—— 原译注)增长指数式快地变小.特别地,任何有限体积三维双曲流形都微分同胚于一个紧三维流形的内部,这个紧三维流形的每个边界分支都是二维环面,其基本群到这个流形的基本群的自然同态是单的.

其次我们有带乘积度量的齐性三维流形:

(1)$S^2\times\mathbf{R}$.

(2)$H^2\times\mathbf{R}$.

以 $S^2\times\mathbf{R}$ 为模型的有限体积局部齐性流形自动紧,且等距于 $S^2\times S^1$ 或 $RP^3\#RP^3$.以 $H^2\times R$ 为模型的有限体积局部齐性流形或者形如$\Sigma\times S^1$,这里 Σ 是有限面积双曲曲面,或者被这样的流形有限覆盖.在后一情形,这个流形是有限面积的二维双曲 orbifold(轨形)上的 Seifert 纤维空间.双曲底空间上的 Seifert 纤维化的例子可以非紧,但它们微分同胚于每个边界分支都是环面的紧流形的内部.同样,每个边界分支的基本群到这个流形的基本群的自然同态是单的.

最后,我们有齐性流形(M,g),这里 M 是一个单连通 Lie 群,9 是一个左不变度量.这个群通过左乘等距地作用在自身上,从而嵌入到 M 的等距群中.这类容许有限体积的局部齐性例子的三维模型有:

(1) 由 3 阶严格上三角矩阵①构成的幂单群(Heisenberg 群)N^3.以这个群为模型的局部齐性流形称为幂零流形.任何有限体积幂零三维流形都被二维环面上的一个圆周丛有限覆盖,因而,特别的,紧.

(2) 可解群可写为一个半直积 $\mathbf{R}^2\times\mathbf{R}^*$ 这里 $t\in\mathbf{R}^*$,在 $\mathbf{R}^2$ 上的作用是对角的,由矩阵

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}$$

① 此处指沿着主对角线的元素都是 1,在主对角线下面的元素都是 0 的矩阵.—— 原译注





给出.以这个群为模型的局部齐性流形称为可解流形.有限体积的可解流形是圆周上的二维环面丛或被这样的流形限覆盖.所有这些例子也都是紧的.

(3) $G = \widetilde{PSL_2}(\mathbf{R})$, $PSL_2(\mathbf{R})$ 的万有覆盖群.这个流形也可视为 H^2 的单位切丛的万有覆盖空间(带有诱导度量).以这个例子为模型的有限体积局部齐性流形是双曲曲面上的圆周丛,或者更一般地,双曲 orbifolds 上的 Seifert 纤维丛.这些例子可以非紧,但若不紧则它们微分同胚于一个所有边界分支都是环面的紧流形的内部,这些环面的基本群到这个流形的基本群的自然同态是单的.

总结所有这些例子共同的主要特征,我们有

推论 1.2.1 设 X^3 连通,可定向,且容许一个有限体积的局部齐性 Riemann 度量,则 X 微分同胚于一个所有边界分支都是环面的紧带边三维流形的内部.进一步,这些环面中的每一个的基本群到 X 的基本群的自然同态都是单的.如果 X 非紧,则它以 H^3, $H^2 \times \mathbf{R}$,或 $\widetilde{PSL_2}(\mathbf{R})$ 为模型,因而 X 或者是三维双曲流形,或者是二维双曲 orbifold 上的 Seifert 纤维空间.

对 Seifert 纤维化三维流形来说,其逆对我们也重要.

定理 1.2.2 设 X 是一个(带边)Seifert 纤维化三维流形,其 Seifert 纤维化的底空间是一个紧二维 orbifold,可能带边.如果底 orbifold 的 orbifold Euler 示性数小于 0,则 X 的内部容许一个以 $H^2 \times \mathbf{R}$ 或 $\widetilde{PSL_2}(\mathbf{R})$ 为模型的有限体积局部齐性度量.如果底空间的 Euler 示性数为 0,则或者 X 容许一个以 $\mathbf{R}^3$ 或 N^3 为模型的局部齐性度量,或者 X 微分同胚于 $T^2 \times I$ 或 Klein 瓶上的扭 I – 丛.如果底空间的 Euler 示性数为正,则或者 X 容许一个以 $S^2 \times \mathbf{R}$ 或 S^3 为模型的局部齐性度量,或者 X 是实心环面.

推论 1.2.3 设 X 是一个紧二维 orbifold 上的 Seifert 纤维化三维流形 $\overline{X}$(可能带边)的内部,则 X 容许一个有限体积的局部齐性度量.除非 $\overline{X}$ 微分同胚于 $T^2 \times I$, Klein 瓶上的扭 I – 丛或者实心环面.

2 三维流形的拓扑

由于几个原因,与二维时不一样,并非每个紧三维流形都容许一个有限体积的(局部)齐性度量.最显而易见的原因与以下事实有关:三维流形不一定是素的(prime),而除了一个例子外,所有局部齐性三维流形都是素的.

2.1 素分解

让我们回顾三维流形的素分解.

定义 2.1.1 给定连通定向光滑 n 维流形 X 和 Y,我们定义它们的连通和

如下.从 X 和 Y 各去掉一个含于它们内部的紧 n 维球体的内部.结果得到带有特别的 $n-1$ 维球面边界分支的两个流形,记为 X_0 和 Y_0.然后把这些球面粘起来,使得 X_0 和 Y_0 的定向沿着公共的 $n-1$ 维球面相配.得到的流形容许一个与 X_0 和 Y_0 的光滑结构相容的光滑结构,而且这个光滑结构在不计微分同胚的意义下不依赖于 n 维球体的嵌入和黏合同胚的选取,它称为 X 和 Y 的连通和,见图 1.注意 X 与一个 n 维球面的连通和微分同胚于 X.三维流形 X 称为是素的.如果它不微分同胚于 S^3,且若球面 $S^2\subset X$ 把 X 分成两块,则其中必有一块微分同胚于三维球体.等价地,每当我们把 X 写成两个流形的连通和时,这两个流形中必恰有一个是三维球面.

三维流形的拓扑中最早的定理之一属于 Kneser(1929)①.

定理 2.1.2 任何闭定向三维流形都能分解成定向素三维流形(称为它的素因子)的连通和.这个分解在不计素因子的次序(和这些因子的保定向微分同胚)的意义下是唯一的.只有可数无穷多个互不微分同胚的素三维流形.

一个本质的(Essential) $S^2\subset X$ 就是在 X 中不界定三维球体的二维球面.如果这个球面分离(X),则沿着它切割 X,然后用三维球体填充得到的洞,就达成了 X 的一个连通和分解.如果另一方面一个球面 $S^2\subset X$ 不分解,则这个过程的效果是从 X 中去掉一个微分同胚于 $S^2\times S^1$ 的素因子.

由于素分解(的存在),所有关于闭、定向三维流形的拓扑的问题都可约化为关于素三维流形的问题.例如:

推论 2.1.3 Σ^3 是同伦三维球面当且仅当它的所有素因子都是同伦三维球面.这样,Poincaré 猜想存在反例,当且仅当它在素三维流形中存在反例.

由球面定理,若 $\pi_2(M)\neq 0$,则 M 含有一个本质的 S^2.其逆等价于 Poincaré 猜想.

2.2 Thurston 几何化猜想

除了以 $S^2\times\mathbf{R}$ 为模型的局部齐性流形外,所有局部齐性三维流形的万有覆盖都微分同胚于 $\mathbf{R}^3$ 或 S^3.易见,如果在一个三维流形的万有覆盖里的每个嵌入二维球面都界定一个三维球体,则在原来的流形里亦然.这样,仅有的具有本质的 S^2 的局部齐性流形就是那些其万有覆盖是 $S^2\times\mathbf{R}$ 的流形.这些是 $S^2\times S^1$ 和 $RP^3\neq RP^3$,后者是非素但容许局部齐性度量的流形的唯一例子.因此,如果我们对赋予局部齐性度量有兴趣,就必须把我们的注意力集中于素三维流形.

Thurston 的几何化猜想是关于素流形的.

猜想 2.2.1(Thurston 几何化猜想) 设 M 是闭、定向、素三维流形,则存在

① 唯一性部分应归功于 Milnor.





二维环面和 Klein 瓶的不交并$\amalg_i T_i^2$ 到 M 的嵌入,使得它的余集的每个分支都容许一个有限体积的局部齐性 Riemann 度量.

这个切割过程的各种分支上的局部齐性度量可以以不同的齐性流形为模型.

由于我们考虑的是可定向流形,一个 Klein 瓶的一个邻域的边界是二维环面,于是得到的几何块是边界分支全是环面的紧流形的内部.注意,一般说来,在不计同痕的意义下,使得几何化猜想的结论成立的环面和 Klein 瓶的这样的族一般可以多于 1 个.例如,设 Σ_2 是亏格为 2 的曲面,则 $\Sigma_2 \times S^1$ 有一个局部齐性度量.设 $T^2 \subseteq \Sigma_2 \times S^1$ 是位于 Σ_2 中一个把它分离成两个穿孔环面的圈之上的环面.则 $\Sigma_2 \times S^1 \setminus T^2$ 的两个分支中的每一个也有一个体积有限的局部齐性度量.为了得到唯一性,我们取 X 中满足结论的分支数最少的族.有了这个条件,在不计同痕的意义下,此族是唯一的.

一个三维流形中的一个亏格至少为 1 的闭曲面称为不可压缩的,如果它的基本群到这个三维流形的基本群的自然同态是单的.注意由推论 1.2.1(在猜想 2.2.1 中),每个环面边界分支的基本群到它界定的每个三维流形的基本群的自然同态都是单的.由 Van Kampen 定理每个环面的基本群到 $\pi_1(M)$ 的自然同态也是单的.这就是说,这些环面是不可压缩的,由类似的推理这些 Klein 瓶也是不可压缩的.这样,我们可把关于素三维流形 M 的 Thurston 几何化猜想叙述为:存在 M 的一个沿不可压缩环面和 Klein 瓶的分解,使得每块的内部都容许有限体积的局部齐性度量.关于一般的紧、可定向三维流形的几何化猜想是说它的每个素因子都满足猜想 2.2.1.

2.3 与 Poincaré 猜想和球面空间型问题的关系

容易从 Thurston 几何化猜想导出 Poincaré 猜想:设有一个满足几何化猜想结论的素同伦三维球面 Σ.由于 $\pi_1(E) = \{1\}$,Σ 不含不可压缩的环面和不可压缩的 Klein 瓶,因此,猜想 2.2.1 中所说的分解对 Σ 来说是平凡的.也就是说 Σ 有一个局部齐性度量.同样,由于 $\pi_1(\Sigma)$ 平凡,Σ 的齐性模型是 Σ 自己(它是紧的).唯一的紧三维模型是 S^3,从而 Σ 微分同胚于 S^3.注意这个论证(除了最后一步)对基本群有限的素三维流形同样适用.结论是对这些流形的几何化猜想蕴涵着它们是三维球面空间型.这些已被 Hopf 分类.

3 Hamilton 的瑞奇流方程

现在让我们引入 Hamilton 的瑞奇流.这是一个流形上的 Riemann 度量的抛物型发展方程.这个方程的一般形式是

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

$$g'(t) = F(g(t))$$

这里,一个解是流形上的度量的单参数族,其时间导数由泛函 $F(g(t))$ 给出. 这个泛函必须与度量有同样的张量型,即对称二阶反变(Contravariant)张量,且应该(我们希望)不牵涉度量的超过二阶的导数. 它还必须是自然的(即不依赖于我们用来局部地写出这个方程的坐标系). 这意味着 F 必须是度量$g(t)$ 自身和下面将要描述的度量的瑞奇曲率 $\mathrm{Ric}(g(t))$ 的线性组合. Hamilton 引入的方程是

$$g'(t) = -2\mathrm{Ric}(g(t)) \qquad ①$$

瑞奇流方程.

这个方程有个尺度不变性:给定度量的一个流 $g(t)$,由

$$h(t) = \lambda^2 g(\lambda^{-2}t) \qquad ②$$

定义一个新的流. 容易看到,如果原来的度量族满足瑞奇流方程,则新的亦然.

显然瑞奇流方程的不动点是瑞奇平坦流形(即瑞奇曲率为 0 的流形). 任何瑞奇平坦三维流形事实上是平坦的. 更一般地,任何 Einstein 流形(由定义 Einstein 流形的瑞奇曲率张量是度量张量的一个倍数)在瑞奇流下是自相似的;当这个倍数为负时,该族是自相似流形的扩张族,当倍数为正时收缩. 而任何 Einstein 三维流形都有常截面曲率. 这样,我们看到三维瑞奇流的不动点(不计共形因子)与有(局部)齐性几何的三维流形之间的密切关系. 这个关系在四维或更高维时不再成立:在那些维数,一个流形是 Einstein 流形并不蕴涵它是局部齐性的.

3.1 曲率

回忆:给定一个 Riemann 度量,在流形的切丛上存在唯一对称联络①

$$\nabla: \text{向量场} \otimes_R \text{向量场} \to \text{向量场}$$

联络通常写为

$$\nabla(X \otimes Y) = \nabla_X(Y)$$

通过定义

$$R(X,Y) = \nabla_X \circ \nabla_Y - \nabla_Y \circ \nabla_X - \nabla_{[X,Y]}$$

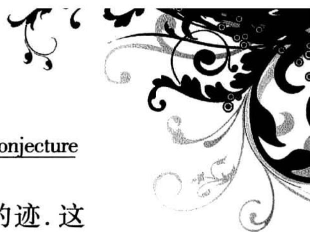
引入曲率算子,它是余切丛的 2 次外幂和切丛的正交同态丛张量积的一个截面. 然后我们定义 Riemann 曲率张量

$$\mathrm{Rm}(X,Y,Z,W) = -\langle R(X,Y)Z, W\rangle$$

这是余切丛的 4 次张量幂的一个截面. 它关于前两个变量和后两个变量都是反对称的,而在交换变量(1,2) 和变量(3,4) 下对称. 这样,我们可视 Rm 为余切丛

① 通常对称联络指无挠联络,须加上与度量相容的条件才有唯一性.





的 2 次外幂上的对称二阶张量. 瑞奇曲率是 Rm 在中间的两个变量上的迹. 这样,若 $e_1,\cdots,e_n$ 是一点处的正交标架,我们有

$$\mathrm{Ric}(e_i,e_j) = \sum_{k=1}^{n}\mathrm{Rm}(e_i,e_k,e_k,e_i)$$

这时流形上的对称二阶反变张量. 在局部坐标系中,它的首项是描述度量的矩阵 g_{ij} 的元素的三阶偏导数的一个线性表达式.

3.2 瑞奇流方程与热方程的关系

我们应视瑞奇流方程为热方程在两方面的推广. 第一,它是关于张量(对称二阶张量)而不是标量函数的. 第二,它是非线性抛物方程,其首项是张量的"热方程". 许多对瑞奇流的理解是通过与关于热方程的结论作类比得到的. 为了指示瑞奇流方程与热方程的非线性形式的类似性,让我们写下数量曲率 R(由定义它是瑞奇曲率的迹)的演化方程. 在三维,这个方程是

$$R'(t) = \Delta R(t) + \frac{2}{3}R^2(t) + |\mathrm{Ric}^0(t)|^2$$

这里 Ric^0 是瑞奇曲率的零迹(Traceless)部分. 从这个表达式可以清楚地看到瑞奇流方程中的负号是被迫加的 —— 没有它我们将得到反向热方程的非线性形式,而这是不适定方程. 瑞奇流中的因子 2 只是为了方便;它可以通过尺度变换去掉.

3.3 Hamilton 的最初结果

中 Hamilton 证明:

(1) 短时存在和唯一性. 如果 g_0 是一个紧流形上的一个光滑度量,则存在依赖于 g_0 的 $\varepsilon > 0$ 和瑞奇流方程的唯一解,它对 $t \in [0,\varepsilon)$ 有定义,且 $g(0) = g_0$.

(2) 奇点形成的曲率刻画. 如果解在时间区间 $[0,T)$ 存在,但不能扩充到任何严格更大的时间区间,则存在流形中的一点 x,使得当 t 趋于 T 时,度量 $g(t)$ 的曲率张量 $\mathrm{Rm}(x,t)$ 无界.

3.3.1 非负瑞奇曲率情形

Hamilton 还完整地分析了瑞奇曲率非负的三维流形上的瑞奇流.

定理 3.3.1(Hamilton) 设 X^3 是具有非负瑞奇曲率的紧、连通三维流形,则下列情形之一发生:

(1) 对所有充分小的 $t > 0$ 瑞奇曲率变得严格正. 在这个情形,瑞奇流在有限时间内产生奇点. 当奇点产生时,流形的直径趋于 0. 对这个度量的演化族进

行尺度变换使得它们的直径是 1,导致一族度量光滑收敛到一个常正曲率度量.特别地,这个流形微分同胚于一个球面空间型.

(2) 存在这个 Riemann 流形的一个有限覆盖,它在诱导度量下是一个正曲率的紧曲面和 S^1 在度量的意义下的乘积.这对该瑞奇流中的所有 Riemann 度量也对.该瑞奇流在有限时间内产生奇点,且考虑的流形微分同胚于 $S^2 \times S^1$ 或 $RP^3 \neq RP^3$.

(3) 度量平坦而演化方程恒定.在此情形,当然,这个流形被 T^3 有限覆盖.特别地,所有这些流形满足 Thurston 几何化猜想的结论.

3.3.2 当 $t \to \infty$ 时的极限

在另一个极端,Hamilton 分析了在一定的额外条件下,当 $t \to \infty$ 时 Riemann 度量会怎样变化. Hamilton 能够证明:

定理 3.3.2(Hamilton) 设一个紧、连通三维 Riemann 流形(M, g_0)上的瑞奇流对所有 $t \in [0, \infty)$ 存在,且规范化的曲率 $t \cdot \mathrm{Rm}(x, t)$ 当 $t \to \infty$ 时有界,则下面两个情形之一发生:

(1) 该流形容许一个平坦度量.

(2) 存在有限体积的完备双曲流形 H_i 的有限集合,以及对每个 $t \gg 1$,一个嵌入 $\varphi_t: \amalg_i H_i \to M$,满足下列性质.尺度变换后度量的拉回 $\varphi_t^*(t^{-1}g(t))$,在 $\amalg_i H_i$ 的每个紧子集上一致收敛于一个常负曲率度量,即双曲度量的伸缩形式. H_i 中平行于无穷远的环面被映到 M 中的不可压缩环面. φ_t 的象的余集有如下结构:它的所有素因子是边界环面不可压缩的 Seifert 纤维化或可解流形.

Hamilton 用来证明 φ_t 的象的余集是一个其素因子都是 Seifert 纤维化(和可解流形 —— 译注)的流形的方法是通过具有有界曲率的塌缩度量的理论. Hamilton 证明当 $t \to \infty$ 时 φ_t 的象的余集上的度量经过t^{-1}的伸缩后塌缩而曲率有界. Cheeger-Gromov 研究了有这个性质的流形,证明它们有所谓 F – 结构.这意味着这个流形局部有正维数的环面的相容作用,于是容易用拓扑技巧对具有 F – 结构的三维流形进行分类.结果是,它们是 Seifert 纤维化流形和可解流形的连通和.

推论 3.3.3 如果以(M, g_0)为初始条件的瑞奇流满足前面定理的假设,则 M 满足 Thurston 几何化猜想.

证明 在定理中所说的第 1 种可能性下这是显然的.设第 2 种成立.(M 中)双曲块并的余集的素分解事实上导致流形 M 的素分解.这样,M 的每个素块都沿着不可压缩环面和 Klein 瓶分解成了双曲块或 Seifert 纤维化的块.由于这些环面和 Klein 瓶不可压缩,从定理1.2.2得知所有 Seifert 纤维化的块或者容许有限体积的局部齐性度量,或者它们自身同构于 $T^2 \times I$ 或 Klein 瓶上的





扭I－丛. 对后两种类型中的每一块,我们把它的边界环面换成中心的环面或Klein瓶.作了这些修改后,我们发现 M 的所有素因子都满足 Thurston 的几何化猜想的结论.

4 Perelman的断言

在过去的两年里投往数学档案馆(网站)的一系列3篇预印本里,Perelman 给出了关于瑞奇流及其与 Thurston 几何化猜想的关系的一系列断言的论据.以下是其中的总体想法.

有限时间奇点在流形的两类区域里发生.第1类由流形的其度量以一种可以控制的方式收缩的分支(例如,正瑞奇曲率分支收缩到一点)组成.第2类中的每个区域是一个细长管子,微分同胚于 S^2 乘以一个区间,或者这样的细长管子并上一个末端具有正曲率的帽子.Perelman 的手术过程是,来到奇异时刻,去掉所有第1类区域(它们是流形的分支),在细长管子的"大"端附近做手术,给它们加上有标准(或几乎标准)度量的球体帽子.这些手术的拓扑效应是去掉一些已知满足 Thurston 几何化猜想的分支,并且在别的分支上作手术,达成一个连通和分解(它可能是平凡的,因为手术产生的一块可能是 S^3).从这里容易推出,如果手术后的流形满足 Thurston 的几何化猜想,则手术前的流形也满足.然后我们取手术后的流形为初始条件延拓瑞奇流.这个过程可无限重复下去,产生带手术的流,它对所有正的时间有定义,且在任意有限时间区间内只有有限多次手术.最后,Perelman断言,Hamilton关于 $t\to\infty$ 时极限性质的结论(定理3.3.2),在带手术的瑞奇流情形的类似物也对,即使没有规范化曲率的假设.由于充分大时刻的流形满足 Thurston 几何化猜想,回溯穿过手术,我们看到初始流形也满足猜想.这样,Perelman 断言建立了 Thurston 几何化猜想,从而也证明了 Poincaré 猜想.

如果我们只对 Poincaré 猜想或更一般地,基本群有限的流形有兴趣,则按照 Perelman 事情得到简化,因为此时带手术的瑞奇流在有限时间后将不复存在,因而不需要定理3.3.2的类似物.

仅为了解说简单起见,我们假设今后我们的三维流形不容许嵌入的射影平面.对拓扑的结果来说这不是严重的限制,因为在一个可定向(三维)流形里,一个嵌入的 RP^2 的一个邻域是 $RP^3\setminus D^3$,因而这个流形存在一个连通和分解分裂出一个 RP^3;在分裂出有限多个这样的 RP^3 后,得到的流形不再含嵌入的 RP^2.

4.1 手术参数

在做手术时要固定3个参数.第一个是粗控参数 $\varepsilon>0$,充分小,是一个万

有常数,一旦取好就不再变.其他两个参数是 t 的非增函数,当 $t \to \infty$ 时趋于0.它们是细控参数 $\delta(t) > 0$ 和尺度参数 $h(t)$.我们再固定一个辅助参数 $\rho(t) = \varepsilon \cdot \delta(t)$.当我们在时刻 t 做手术时.我们利用 ε 和这些参数在时刻 t 的值以便做手术.Perelman 的断言是,如果恰当地选取这些常数和函数,则下面的结论就成立.现在它们被看成是固定的.

4.2 瑞奇流中的大曲率区域

为了做手术我们需要了解瑞奇流的数量曲率大的区域.当然.由于在任何流中都可能对度量和时间进行尺度变换,我们必须通过某种方式规范化以得到大曲率的不变概念.这样,我们调整一下使我们的流有规范化的初始条件,即在 $t = 0$ 时在每一点 Riemann 曲率的绝对值都不超过1,且任何半径为1的度量球体的体积至少是 $\mathbf{R}^3$ 中单位球体的体积的一半.这可以通过像在方程②中那样对度量和时间用互为倒数的正常数进行尺度变换达到.从现在起,我们隐含地假定所考虑的流都有规范化初始条件.

定义 4.2.1 三维 Riemann 流形 (X, g) 中以 $x \in X$ 为中心的 ε - 瓶颈是 X 中的一个区域 T,使得 $x \in T$ 且有一个微分同胚 $\varphi: T \to S^2 \times [-\varepsilon^{-1}, \varepsilon^{-1}]$ 把 x 送进 $S^2 \times \{0\}$,并且 $R(x)$ 倍的 g 在 φ 下的推前(Pushforward)与 S^2 上 Gauss 曲率为1/2的度量和 $[-\varepsilon^{-1}, \varepsilon^{-1}]$ 上的通常度量的乘积在 C^N - 拓扑下 ε - 接近,见图1.

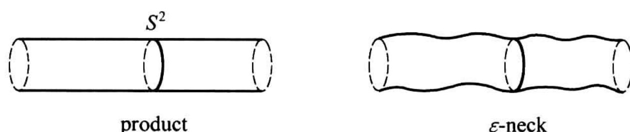


图1

断言 4.2.2(Perelman) 存在 $r > 0$ 使得下述事情成立,设 $(M, g(t))$ 是满足规范化初始条件的瑞奇流,它对 $0 \leqslant t < T$ 有定义,这里 M 是闭、可定向三维流形.则对瑞奇流中满足 $R(x, t) \geqslant r^{-2}$ 的任何点 (x, t),下列情形之一成立:

(1) M_t 的含 (x, t) 的分支微分同胚于球面空间型.

(2) (x, t) 是 M_t 中的一个 ε - 瓶颈的中心.

(3) (x, t) 含在 M_t 的一个微分同胚于 D^3 的子流形里,这个子流形的边界里的所有点都是 M_t 中的一个 ε - 瓶颈的中心.(这样的区域称为 ε - 帽子,见图2)

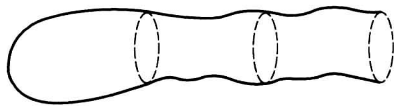


图2 ε - 帽子



4.3 奇异时刻的大曲率区域

现在设当 $t \to T$ 时瑞奇流产生奇点.上面的定理在这个时刻有一个扩充.存在一个开子集 $\Omega_T \subset M$,它由 Riemann 曲率张量保持有界的所有点组成.由 Hamilton 和施(施皖雄——原译注)的结果,当 $t \to T$ 时在这个开子集上存在一个极限度量.我们记之为 $(\Omega_T, g(T))$.当然,下面是断言 4.2.2 的直接推论.

断言 4.3.1 存在 $r > 0$ 使得下面所述成立.设 $(M, g(t))$ 是满足规范化初始条件的瑞奇流,对 $0 \leqslant t \leqslant T$ 有定义,且在 $t = T$ 时刻产生奇点,这里 M 是一个闭、可定向三维流形,则对任何满足 $R(x, T) \geqslant r^{-2}$ 的点 $(x, T) \in \Omega_T$,下列事情之一成立:

(1)Ω_T 的含 (x, T) 的分支微分同胚于一个球面空间型.

(2)(x, T) 是 Ω_T 里的一个 ε - 瓶颈的中心.

(3)(x, T) 含于 Ω_T 里的一个 ε - 帽子.

事实上我们可以给出关于 Ω_T 的端的更精确的信息.

断言 4.3.2 我们固定 $r > 0$ 和一个手术参数 $\delta > 0$.令 $\rho = r \cdot \delta$.设 $\Omega_T(\rho)$ 是 Ω_T 中满足 $R(x, T) \leqslant \rho^{-2}$ 的点 (x, T) 所成的子集.(图 3)

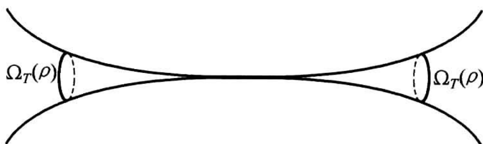


图 3

(1)Ω_T 中的一个 ε - 管是一个微分同胚于 $S^2 \times I$ 的子流形,其每一点都是 Ω_T 中的一个 ε - 瓶颈的中心.

(2) 一个 ε - 回路是 Ω_T 的一个分支,它是一个闭流形,其中每一点都是一个 ε - 瓶颈的中心.这样的流形自动微分同胚于 $S^2 \times S^1$.

(3) 一个 ε - 角(关于 r 和 δ)是 Ω_T 中一个微分同胚于 $S^2 \times [0,1)$ 的闭子集 H,其边界含于 $\Omega_T(\rho)$,且 H 的每一点都是 Ω_T 中的一个 ε - 瓶颈的中心.当然,在 H 的另一端数量曲率趋于无穷.

(4) 一个双 ε - 角是 Ω_T 的一个微分同胚于 $S^2 \times (0,1)$ 的分支,其每点都是 Ω_T 中的一个 ε - 瓶颈的中心.在这个分支的两端数量曲率都趋于无穷.

(5) 一个加帽 ε - 角是 Ω_T 的一个微分同胚于 $\mathrm{int}(D^3)$ 的分支,其每点或者是一个 ε - 瓶颈的中心,或者含于一个 ε - 帽子.在加帽 ε - 角的末端附近数量曲率趋于无穷.

断言 4.3.3(Perelman) 存在 $r > 0, \delta > 0$ 和 $\rho = r \cdot \delta$,使得对于任何满

足规范化初始条件且在 $t = T$ 时刻产生奇点的瑞奇流 $(M, g(t))$(M 是一个闭、可定向三维流形),下列事情成立.Ω_T 中任何满足 $R(x, T) \geqslant r^{-2}$ 的点 (x, T) 都含于下列区域之一:

(1)Ω_T 的一个微分同胚于球面空间型的分支.

(2)一个微分同胚于 $S^2 \times S^1$ 的 ε - 回路.

(3)一个边界分支在 $\Omega_T(\rho)$ 中的 ε - 管.

(4)一个边界在 $\Omega_T(\rho)$ 中的 ε - 帽.

(5)一个边界在 $\Omega_T(\rho)$ 中的 ε - 角.

(6)一个加帽 ε - 角.

(7)一个双 ε - 角.

为了处理在管子和帽子中产生的奇点,Perelman 引入沿着一个二维球面的手术.(在研究某种四维瑞奇流时 Hamilton 更早引入了沿着一个三维球面做手术的类似思想)这就要求引入 Perelman 所说的标准解.

4.3.1 标准解

标准解是初始条件为 $(\mathbf{R}^3, g_0)$ 的瑞奇流,这里度量 g_0 是一个 C^∞ 度量,它是 $t = 0$ S^3 的上半球面上的一个接近于常正曲率 1/2 的度量与 $S^2 \times [0, \infty)$ 上的乘积度量(其中 S^2 上的度量有常曲率 1/2)的并.我们选取这个度量使得那个明显的 $SO(3)$ 对称是一个等距作用.这个作用的唯一的不动点 $(p, 0)$ 称为这个标准解的中心.结果,尽管初始 Riemann 流形非紧,带有这些初始条件的瑞奇流 $(\mathbf{R}^3, g(t))$ 是唯一的,且对所有 $t \in [0,1)$ 有定义.对我们的讨论来说,这个流的重要性质有:对每个 $t \in (0,1)$,度量 $g(t)$ 是完备的且有有界的正截面曲率,在无穷远渐近于 S^2 上曲率为 $1/2(1-t)$ 的度量与实直线上标准度量的乘积,而且那个 $SO(3)$ 作用对每个 $g(t)(t < 1)$ 都是等距.(图 4)

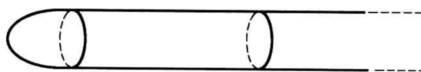


图 4　标准解

4.3.2 手术

现在想法是,从满足规范化初始条件的瑞奇流开始,行进到奇异时刻 T.我们去掉 Ω_T 的所有这样的分支,它们是属于断言 4.3.3 中前两类的 M 的分支.我们也去掉 Ω_T 的所有加帽 ε - 角和双 ε - 角分支.现在在每个 ε - 角里,可以找到一点 (x, T),它是一个半径为 $h(T)$ 的 $\delta(T)$ - 瓶颈的中心.我们沿着这个 $\delta(T)$ - 瓶颈的中心二维球面把 Ω_T 切开,并去掉含着 ε - 角的末端的那块.然后对新得到的边界.S^2 用一个来自标准解的初始条件的半径为 λ 的球体去粘,这



里 λ 是某个固定的万有常数.我们用 $h(T)^2$ 对后面这个球体的度量作尺度变换,使得我们要把它们粘在一块的二维球面上的度量彼此接近.我们利用单位分解在这个并上造一个光滑度量.(图 5)

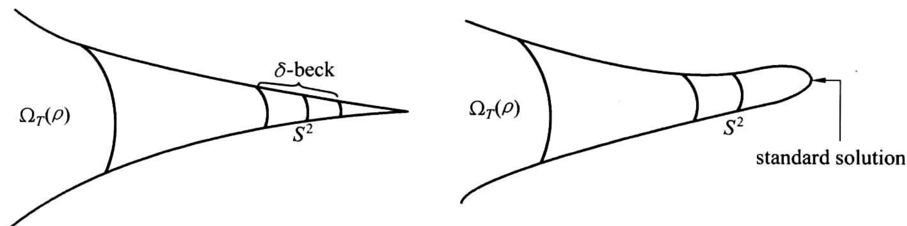


图 5

流形$(M_T, g(T))$的绝大部分都是时刻 T 瑞奇流的结果.但是我们已经走进在这个奇异时刻产生的奇点的邻域并去掉它,同时缝上三维球体上我们专门造的固定 Riemann 度量.这样造出闭 Riemann 流形 M_T 后,我们利用它作为瑞奇流在时刻 $t = T$ 的初始条件.这样我们延拓带手术的瑞奇流到下一个奇异时刻.在下一个奇异时刻 T',我们做完全同样的手术过程,然后重新开始瑞奇流.这样 Perelman 造出了一个以任何紧三维 Riemann 流形为初始条件的带手术的瑞奇流.这个带手术的流对所有 $t \in [0, \infty)$ 都存在.

在手术过程下,三维流形改变拓扑类型.拓扑变化的类型有 3 重:

(1)(管子手术)从流形里去掉一个管子 $S^2 \times I$,并在每一端粘上三维球体.(图 6).

(2)(帽子手术)从流形里去掉一个三维球体并粘上另一个.(图 7).

(3)(分支塌缩手术)流形的一个分支全部被去掉.

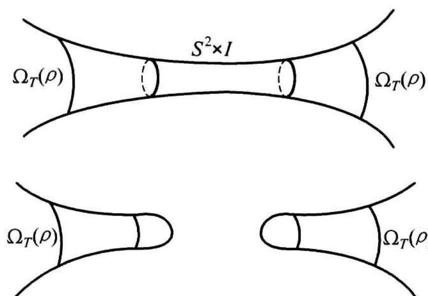


图 6　管子手术

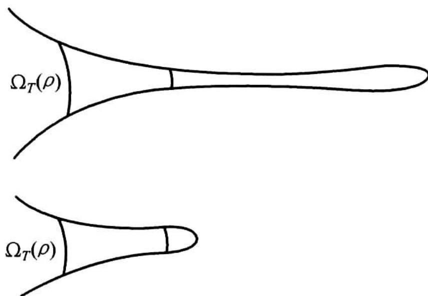


图 7　帽子手术

第一种类型的操作是通常的拓扑手术.这就是几何化猜想中要求的连通和分解是怎样在带手术的瑞奇流中达成的.(当然没有理由期望每次手术都产生一个非平凡的连通和分解.某些会简单地分裂出 S^3,但这在拓扑上无影响)第 2 类操作在度量上有意义,因为所考虑的球体有不同的度量,但在拓扑上它

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

没做什么.最后一类操作可以显著地改变拓扑,因为整个分支都被去掉,但是,如我们所见,这些分支中的每一个都微分同胚于球面空间型或 $S^2 \times S^1$,因而满足几何化猜想的结论.这样,我们在去掉原来的三维流形的素分解中满足猜想的块,而保留那些我们还未知的块.从这个描述容易看到,如果手术过程的结果满足 Thurston 几何化猜想,则初始流形也满足.

因为沿着 S^2 做手术的过程在任意给定时刻 $T_0 < \infty$ 之前从这个三维流形中去掉固定数量的体积(到底多少依赖于参数 $\delta(t)$ 和 $h(t)$ 对 $t \leqslant T_0$ 的下界),又因为在瑞奇流下,体积增长不会比时间的一个固定指数率更快,所以对任意 $T_0 < \infty$,在时刻 T_0 之前管子手术和帽子手术的次数有界,这意味着对所有 $t \leqslant T_0$,在这些手术下造出来的分支的数目有一个一致的界,从而被塌缩分支手术去掉的分支的数目有界.这就是说在时刻 T_0 以前只有有限多次手术.当我们允许时间趋于 ∞ 时,可以有无限多次手术.去掉非 S^3 分支的手术只能有有限多次,沿着本质的 S^2 的管子手术也如此.所有其他手术或者是帽子手术(它们在拓扑上平凡),或引入新的 S^3 的连通和分解,或去掉 S^3 分支的手术.设 (M,g) 是闭三维 Riemann 流形.设 M_t 是以 (M,g) 为初始条件的带手术的瑞奇流在时刻 t 的结果.如果我们能够证明对某个 t,流形 M_t 满足 Thurston 几何化猜想的结论,则 M 也满足这个猜想的结论.

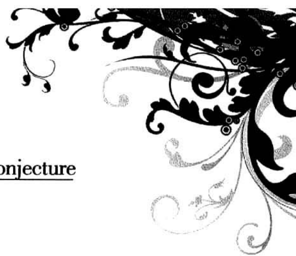
4.3.3 关于手术参数的评论

从上面对手术的描述还不清楚为什么我们需要允许手术参数 $\delta(t)$ 和 $h(t)$ 随 t 增长而减小.让我们对其理由给一个简明的指示.断言 4.2.2 和 4.3.3 中的常数 r,δ,h 依赖于流形中所有小球的体积与其半径立方之比.这个估计首先是对产生于满足规范化初始条件的瑞奇流的流形建立的.人们必须对满足规范化初始条件的带手术的瑞奇流证明类似的估计,结果这些估计更弱且事实上依赖于到那时为止使用的手术参数.(我们不能对从手术时刻 T 开始的瑞奇流用那个(最初的——原译注)结果,因为没有理由期望这个流形是规范化的)从而,必须允许这些结果中的常数 r:依赖于 t 且当 $t \to \infty$ 时趋于 0.这迫使我们允许常数 δ 依赖于时间且当 $t \to \infty$ 时趋于 0.事实上,这些参数是由在越来越大的时间段上的一个复杂的连锁归纳来定义的.

4.4 无穷时的极限

在 Perelman 的论证中下一步是分析当 $t \to \infty$ 时带手术的瑞奇流的极限.在这部分的论证中他再次沿着 Hamilton 开辟的道路前进,尽管现在由于两个原因情形更复杂了:(i) 我们在考虑带手术的瑞奇流,而不是瑞奇流;(ii) 没有关于(规范化的)曲率的界的假设.尽管有这些复杂性 Perelman 断言 Hamilton 的许多





分析都可以修改:

断言4.4.1(Perelman)　设$(M_t, g(t))$是带手术的瑞奇流,对$0 \leqslant t < \infty$有定义.

(1)存在有限体积的完备双曲流形H_i的有限集合和一个嵌入(对所有充分大的t)

$$\varphi_t : \coprod_i H_i \to M_t$$

对充分大的t,φ_t,在$\coprod_i H_i$的任意大的紧子集上任意接近于一个等距,假若我们用ct^{-i}(c是不依赖于t的适当常数)对M_t上的度量作尺度变换.

(2)对所有$t \gg 1$,H_i的平行于无穷远的环面映到M_t中的不可压缩环面.

(3)对所有$t \gg 1$,$\coprod_i H_i$的一个充分大紧子集的象的余集有一个任意塌缩而曲率有下界的度量.

对一个完备Riemann流形中的一点x,定义$r(x)$为满足以下条件的$r \geqslant 0$的上确界:在以x为中心、r为半径的度量球体$B_r(x)$上$\mathrm{Rm}(y) \geqslant -r^2$.说一个度量在一点$x$是$w$-塌缩而曲率有下界的,是指$\mathrm{Vol}(B_r(x)(x)) < wr^3(x)$.一个流形是$w$-塌缩而曲率有下界的,如果对这个流形的每一点$x$上面所述成立.说$\varphi_t$的象的余集当$t \to \infty$是任意塌缩而曲率有下界的,是指给定任意$w > 0$,存在$T(W) < \infty$,使得对所有$t \geqslant T(w)$,$X_t = M_t \setminus \mathrm{Im}(\varphi_t)$是$w$-塌缩而曲率有下界的.

M_t的一些块等距地逼近固定的双曲块的事实蕴涵着所有在充分大的时刻做的手术都是在流形M_t的塌缩区域做的.

由这个断言推出对任何$t \gg 1$我们可以沿着不可压缩环面把M_t分离成若干块:微分同胚于有限体积的完备双曲流形的块$\varphi_i(\coprod_i H_i)$,和度量w-塌缩而曲率有下界的块X_t(这里w依赖于t且当t趋于无穷时趋于0).剩下来要证明对所有$t \gg 1$,X_t的素因子可以沿着不可压缩的环面和Klein瓶分解成容许有限体积的局部齐性几何的块.

4.5　塌缩区域

为了做这个我们引用一些结果,它们来自度量塌缩而曲率有下界的流形的理论.在我们的情形,我们有一族三维流形X_t,它们当$t \to \infty$时塌缩而曲率(局部)有下界.想法是过渡到(X_t的一个子列的)极限X.这个极限是一个比Riemann流形更一般的度量空间,但在其中曲率有下界的概念仍有意义.这些东西的自然范畴是Aleksandrov空间的范畴.这样的空间有一致的维数,且极限的维数不超过这个序列中的流形的维数.

我们把这个一般理论用到带手术的瑞奇流的塌缩区域当$t \to \infty$时产生的极限.除了一个例外,对这些极限发生的事情是,存在不可压缩的环面和Klein

瓶把塌缩区域分成块，每一块有一个一维或二维的 Aleksandrov 空间极限．其 Aleksandrov 空间极限为二维的块是双曲 orbifold 上的 Seifert 纤维化，或者是幂零流形分支．其 Aleksandrov 空间极限为一维的块是这样的分支：它们是可解流形，或幂零流形，或 $T^2 \times I$，或 Klein 瓶上的扭 I – 丛，或 $S^2 \times I$．我们沿着 $S^2 \times I$ 块做手术以把它们从这个单子中去掉．我们也可通过简单地把 $T^2 \times I$ 和 Klein 瓶上的扭 I – 丛的核（Cores）加到我们要沿之切割的环面和 Klein 瓶的集合里去，而从这个单子中去掉这些块．于是所有剩下的块或者是可解流形，幂零流形，或者是二维双曲 orbifolds 上的 Seifert 纤维空间，所有这些流形都有有限体积的局部齐性度量．这样，像以前一样，我们证明对所有 $t \gg 1$，X_t 的每个素因子都可以沿着不可压缩的环面和 Klein 瓶分解，使得每块有有限体积的局部齐性几何．同样的陈述．于是显然对 M_t 成立（对所有 $t \gg 1$）．

没有包含在上面的描述中的例外情形是 M_t 的这样的分支 Y_t，它们不含 M_t 的任何双曲块，且在做尺度变换使其直径为 1 后，当 $t \to \infty$ 时它的体积有正的下界．这样做尺度变换后，当 $t \to \infty$ 时，过渡到这样的例外分支 Y_t 的一个子序列，我们得到一列瑞奇流，光滑收敛到一个极限瑞奇流．由 Hamilton（和 Ivey——原译注）关于曲率夹挤（Pinching）的结果[9]，极限瑞奇流的截面曲率处处非负．进一步，由构造，截面曲率的最小值是 0．引用 Hamilton 关于瑞奇流的强极大值原理，我们得到极限瑞奇流是一个平坦流形上的恒定流．由于做了尺度变换后的 Y_t 光滑收敛到这个极限，我们得到，对所有 $t \gg 1$，Y_t 有平坦度量，因而满足 Thurston 几何化猜想．

这就覆盖了所有情形，从而确立：对所有充分大的 t，Thurston 几何化猜想的结论对 m_t 成立．如我们已指出的，一旦我们知道对 $t \gg 1$ 的结论，我们就可以利用倒推归纳法，穿过手术时刻，对带手术的瑞奇流的初始流形建立这个结论．由于每个闭、可定向拓扑三维流形都有一个规范化的度量，因而可被用为一个满足规范化初始条件的带手术的瑞奇流的初始条件，于是推出 Thurston 几何化猜想对所有闭、可定向三维流形成立．

严格说来，我们假设了（所考虑的）可定向三维流形不含嵌入的 RP^2，但如我们已指出的，任何闭、定向三维流形是有限多个 RP^3 和一个不含嵌入的 RP^2 的流形的连通和．对闭、可定向、不含嵌入的 RP^2 的三维流形建立 Thurston 几何化猜想，就对所有闭、可定向三维流形建立了猜想．这就完成了关于 Perelman 对 Thurston 几何化猜想的处理和断言的概述．

4.6　对 non-acyclic 维流形的有限时间消亡

关于 Poincaré 猜想和球面空间型问题还有些有趣的事要说．我们说一个带手术的瑞奇流在一个有限时刻 T 消亡，如果 T 是一个手术时刻，且在时刻 T 这





个流形全部被手术过程去掉.按照Hamilton的定理3.3.1,如果我们从一个有正瑞奇曲率的流形出发,在第一次手术时刻就是这个情形.Perelman断言,如果初始条件是(M,g)而M的素因子都不是acyclic(acyclic即有一个可缩万有覆盖),则带手术的瑞奇流在某个有限时刻消亡.特别地,如果Σ^3是一个同伦三维球面,则它的所有素因子也都是同伦三维球面,因此它的素因子都不是acyclic,从而带手术的瑞奇流在有限时间内消亡.这意味着为了证明Poincaré猜想,或更一般地对基本群有限的三维流形进行分类,因而解决球面空间型问题,我们不需要上一子节中关于塌缩度量的结果(它们被用于理解带手术的瑞奇流当$t \to \infty$时的拓扑).

5 最后的评论

我强调了Perelman的结果中处理带手术的瑞奇流的方面,理由当然是这些是拓扑应用需要的陈述.但是,如Perelman自己弄清楚的,他作出的最有意义的进展是关于瑞奇流的.我想总结这些.Perelman运用了一个自举的(boot-strapping)论证.首先,他完全了解了三维的梯度收缩解(gradient shrinking solution)——仅有两类:常正曲率流形的收缩族,和柱面$S^2 \times \mathbf{R}$,这里S^2上的度量族是常曲率度量的收缩族(或被这个流形($S^2 \times \mathbf{R}$——原译注)二重覆盖的流形).第二,他能够对非平坦、完备、古老(Ancient)、体积不塌缩、曲率非负的解的空间导出紧性结果(可以相差尺度变换).第三,他证明每个这样的解或者是一族收缩的常正曲率度量,或者其中有越来越大的度量球(当我们回溯时间)非常接近于收缩柱面(或者它在自由对合下的商).为了建立这些结果,Perelman引入了约化长度和约化体积,并对它们证明了基本的单调性结果.由于Hamilton(和Ivey——原译注)的夹挤结果,关于非负曲率古老解的结果与紧三维流形上的瑞奇流有关.利用这个结果,Hamilton证明,瑞奇流中大曲率区域的某种重尺(Rescaling)极限是非平坦、曲率非负且有界的古老解.Perelman从两方面扩充了Hamilton的结果.首先,他证明,围绕所有大曲率区域人们总是可以取这样的重尺极限,而这些重尺极限自动体积非塌缩(这是约化长度函数的另一个应用).因此,所有以大曲率的点为中心的重尺极限适合他的分类规划.这是Perelman对奇点形成(它当曲率趋于无穷时发生)和产生的手术过程的了解的基础.我在这里是对瑞奇流来阐述事情的,而这些结果是为了定义手术过程而用的.然而如Perelman所说,把这些结果从瑞奇流扩充到带手术的瑞奇流,虽然“技术上十分复杂,但不牵涉本质上新的思想”.

Perelman关于带手术的瑞奇流的效果的图像完美地适合Thurston几何化猜想.有限时刻的手术达成连通和分解(成素因子),且去掉non-acyclic的素因子,

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

即那些万有覆盖是 S^3 或 $S^2 \times \mathbf{R}$ 的素因子.然后,在分析当 $t \to \infty$ 的极限时产生沿环面和 Klein 瓶的分解 —— 沿着环面和 Klein 瓶分解成膨胀的双曲块和塌缩块.瑞奇流在双曲块上产生的度量适当重尺化(Rescaled)后收敛于双曲度量,度量塌缩的块是通过 Aleksandrov 空间即曲率有下界的空间的理论来研究的.目前对瑞奇流在这些块上产生的度量的性质的了解还没有 —— 除了塌缩判别准则.然而,我们从瑞奇流和 Aleksandrov 空间得到拓扑的而不是度量的结论.幸运地,这些结论强到足以允许拓扑类型的分类,而所有这些拓扑类型已知有局部齐性度量(基本上是通过对维数作归纳).

一个与在瑞奇流下在塌缩区域中发生了什么相关的未解决问题是:对带有任意初始条件的瑞奇流,对所有正的时间,是否一般只有有限多次手术?





附录Ⅶ　丘成桐先生在晨兴数学中心的演讲

先生们,女士们:

今天我将会告诉你们数学上的一页篇章是如何结束和新的篇章正在开始.

请允许我先从一些基本的观察开始.

几何结构

几何学的主要目的是描述与分类有趣的几何结构.我们在日常生活中看到许多有趣的几何结构.

我举几个例子:



曲面的亏格就是环柄的数目.

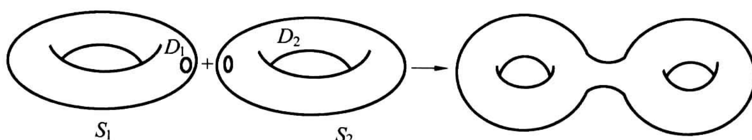
$$\frac{dg_{ij}}{dt}=-2R_{ij}$$

连通和

构造曲面的一个抽象和主要的方法是作曲面的连通和.

连通和 $S_1 \# S_2$ 是通过删除圆盘 D_i,并且钻孔曲面 $S_i - D_i$ 通过微分同胚 $h:\partial D_1 \to \partial D_2$ 粘合起来,于是

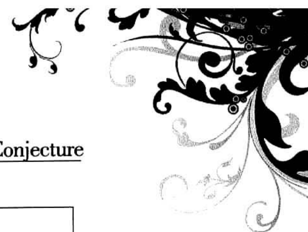
$$S_1 \# S_2 = (S_1 - D_1) \cup_h (S_2 - D_2)$$



曲面结构定理

定理(曲面分类定理)

任意闭的可定向的曲面是如下曲面之一:球面、环面或有限多个环面的连通和.



共形几何

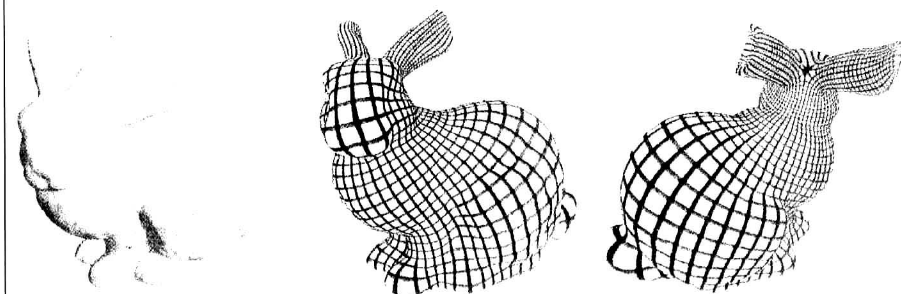
为了更深入理解曲面,庞加莱建议理解这些二维对象上的共形几何.

例子

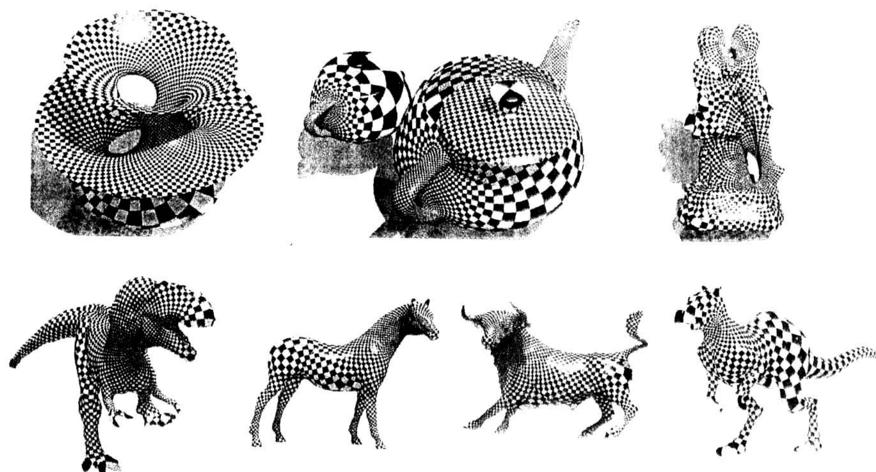
在地球上我们利用经线和纬线来确定方位.它们互相垂直.当我们将方形的地图映到球面上的时候,距离产生了扭曲.比如,北极附近很小的区域在方形地图上是很大的区域.不过,经线与纬线的正交性在映照下保持不变.所以,如果一艘船在海上航行,我们可以用地图精确地指引它的航向.

共形几何

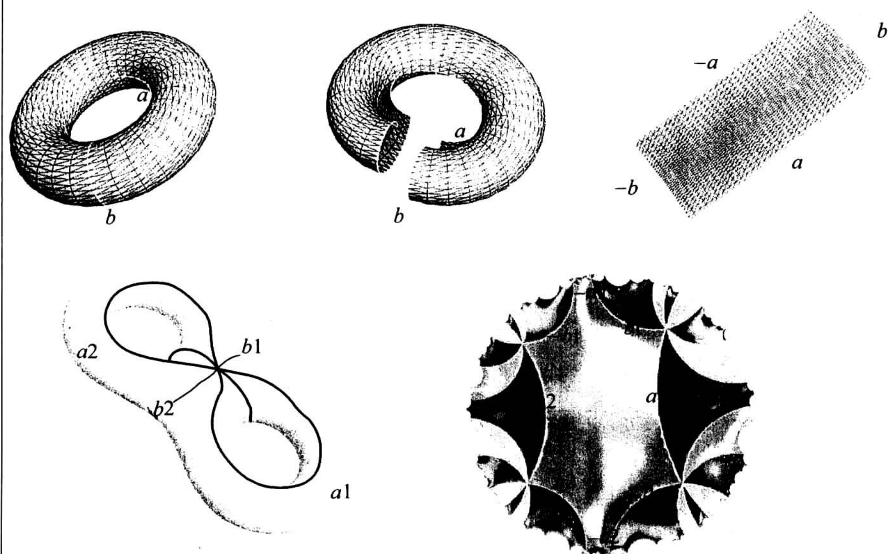
庞加莱(Poincaré)发现,我们可以在任何曲面上绘制经线与纬线.



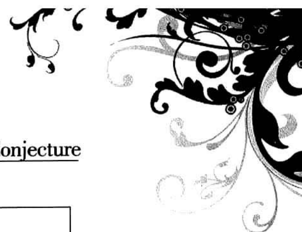
共形结构



我们可以沿着曲面上某些特殊的曲线切割,然后把曲面在平面或圆盘上展开.在这个过程中,经线与纬线保持不变.





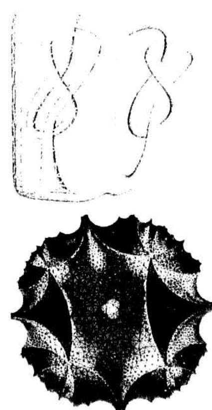


曲面上共形结构的例子

定理(庞加莱单值化定理)

任意二维封闭空间必与一常高斯曲率空间共形等价.





球面　　欧氏　　双曲

曲面上的 Hamilton 方程

我们可以通过曲率变动任意曲面.这种形变就是曲面的 Hamilton 的 Ricci 流.这种形变最后得到常曲率空间.这一方法是 Hamilton 发明的,可用来改变任意维空间.

三维流形

到目前为止,我们所讨论的空间只有两个自由度.与束缚于曲面上的虫子所看到的二维空间不一样,我们所生存的空间有三个自由度.虽然我们的三维空间看起来是平坦的,但还有许多自然而不平坦的三维空间.

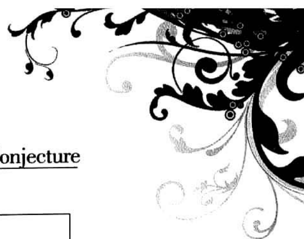
三维流形

例子

相空间

在20世纪初,庞加莱研究粒子动力学的相空间.相空间由(x;v),即粒子的位置与速度组成.例如,如果一个粒子在二维曲面Σ上以单位速度自由移动,那么这个粒子就有三个自由度.这就产生了一个三维空间M.





纤维丛

如果我们对 M 上每个点$(x;v)$,赋以点 $x\in\Sigma$,我们得到一个从 M 到Σ 的映射.当我们固定点 x,v 可以取任意单位向量,因此 v 可以在单位圆上自由移动.我们称 M 是Σ 上的纤维丛,而它的纤维是单位圆.

庞加莱猜想

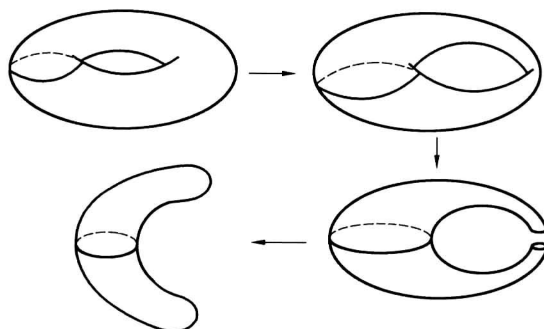
高维拓扑学可以说是从庞加莱的问题开始:

一个闭的三维空间,若其上的每条闭曲线都可以连续收缩到一个点,那么从拓扑上来看,这个空间是否就是球面?

这个问题不仅是一个著名的难题,而且是三维拓扑理论的中心问题.

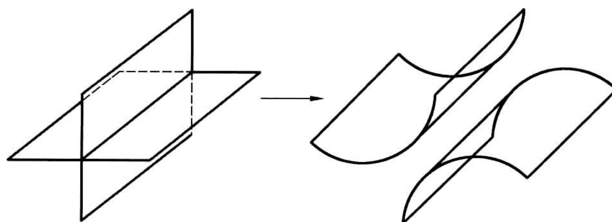
拓扑手术

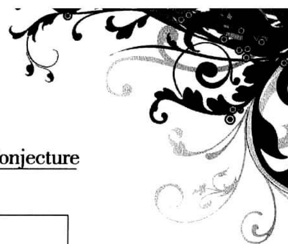
拓扑学家研究这个问题已经有一百多年的历史了.主要的工具是切割和黏合,或称手术,来简化一个空间的拓扑.



拓扑手术

在20世纪70年代以前,主要的工具有Dehn引理,提供了将自相交叉的曲面简化为无交叉曲面的工具.





拓扑手术

定理(Dehn 引理)

如果存在从圆盘到三维空间的一个映射,且不在圆盘边界上自相交叉,那么存在另一个到三维空间的没有自交叉的映射,且限制在边界上与原来的映射相等.

Dehn 引理的一种基于极小曲面理论的版本是 Meeks – 丘成桐发现的,对以后的发展很有帮助.

拓扑手术

第二个工具是 Haken 引入的不可压缩曲面的构造.它被用来将三维流形切割成片.Walhausen 用这一方法证明了重要的定理.(不可压缩曲面是一种嵌入曲面,且具有如下性质:如果一条闭环路不能在曲面上收缩到一个点,那么它也不能在三维空间中收缩到一个点)

特殊曲面

有几个重要的一维和二维空间在理解三维空间的过程中起了重要的作用.

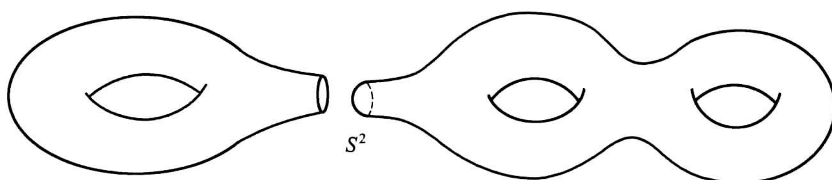
1.圆周

Seifert 构造了许多三维空间,可以写成圆周的连续族.上面提到的相空间是 Seifert 空间的一个例子.

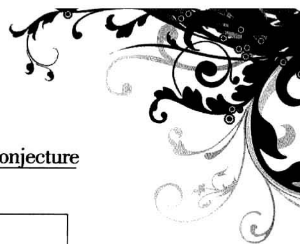
特殊曲面

2.三维球面

我们可以通过在两个三维空间上的各挖去一个实心球,然后沿着球面粘合起来.



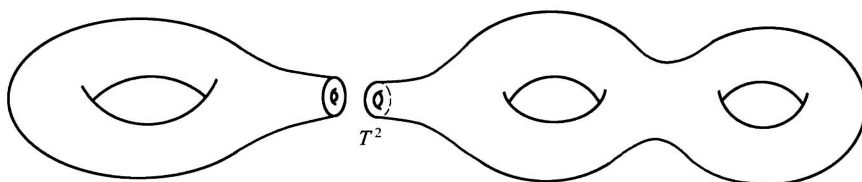
相反,Kneser 和 Milnor 证明每个三维空间可以通过球面唯一分解成不可约分支.一个空间称为是不可约的,如果每个嵌入球面都是这个空间中的一个三维球的边界.



特殊曲面

3.环面

Jaco-Shalen,Johannson 的一个定理说,我们可以通过沿环面切割作进一步分解.



三维空间的结构

几何化猜测(Thurston):

三维空间的结构是由如下的基本空间所合成的:

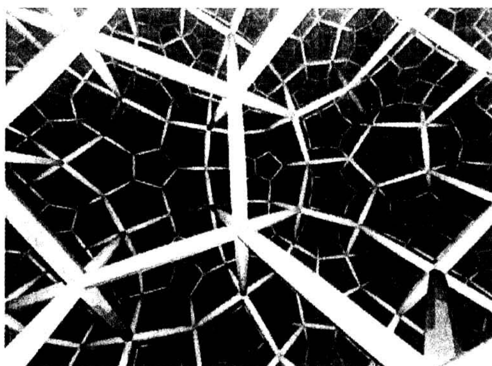
(1)(庞加莱猜想)如果三维空间上每条闭环路都可以收缩到一个点,那么这个空间就是三维球面.

(2)(空间形式问题)将三维球面上的点等同起来得到的空间.这由线性等距的一个有限群所支配,类似于晶体的对称.

(3)Seifert 空间及其类似于(2)用有限群得出的空间.

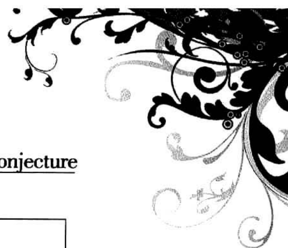
(4)(Thurston 猜测:双曲空间)边界由环面构成的三维空间,空间中每个二维球面都是某个球的边界,每个不可压缩的环面可以用适当的方法形变到边界;这种空间被猜测为带有常负曲率的空间,并且可以通过双曲球的一个离散对称群得到.

三维双曲空间



用十二面体拼成的双曲空间(Charlie Gunn)





三维空间的结构

Thurston 猜测将三维空间的分类简化为群论问题,发展出了许多工具.他和一些后来的学者证明了当三维空间足够大时(这是 Haken 和 Walhausen 所研究的空间),猜想成立.(一个空间里如果有非平凡和不可压缩的嵌入曲面,我们称它为足够大的)

可惜 Thurston 的证明方法很难用到最一般的流形上.

几何分析

另一方面,从 20 世纪 70 年代开始,一群几何分析学家应用非线性偏微分方程来构造空间的几何结构.Yamabe 考虑了将一个空间共形地变为常数量曲率空间.可是这种方法不能用来区分空间的拓扑.

几何分析

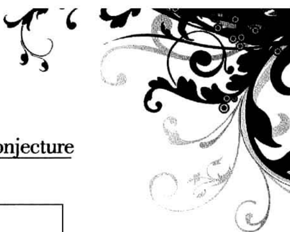
一个重要的发现是 1976 年凯勒–爱因斯坦空间的构造.事实上,我用这个方法证明了复情形的庞加莱猜想.在复几何中被称为 Severi 猜测,即每个同伦等价于复射影平面的复曲面必是复射影平面.

几何分析

将几何与分析的想法结合起来理解几何与拓扑的学科称为几何分析.而这一学科可以追溯到 19 世纪 50 年代,在过去 30 年中有了长足的发展.

这一学科有两大支柱:非线性分析与几何.由于许多学者的努力,这两个学科在 20 世纪 70 年代都变得很成熟.(见我的综述文章 *Perspectives on Geometric Analysis* in Survey in Differential Geometry, Vol, 10, 2006).





爱因斯坦空间

我现在介绍一下几何分析的想法如何用来解决庞加莱猜想.

在三维空间情形,我们需要构造爱因斯坦结构,这是受到了重力理论中的爱因斯坦方程启发.对任何一个三维空间结构,我们找一种方法将它形变到一个满足爱因斯坦方程的空间结构.这种形变必须依赖于空间的曲率.

爱因斯坦方程

爱因斯坦的相对论告诉我们,在重力影响下,时空具有曲率.空间不断地改变.空间的整体拓扑随着曲率(重力)的分布而变化.相反的,整体拓扑非常重要,它提供了重力分布的限制条件,也可以看做重力的源头.

爱因斯坦结构

假设

我们假设三维空间是紧致无边的(也就是闭的).

Ricci 张量

在三维空间中,空间的曲率从不同方向测量会不一样.这种测量受 Ricci 张量 R_{ij}支配.这本质上是空间的物质张量.

Ricci 曲率

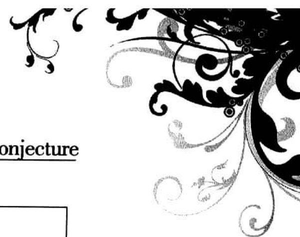
数量曲率

与方向无关的一个重要的量是数量曲率 R.它是 R_{ij} 的迹,可以用来测量测地球的扩张或收缩

$$\mathrm{Volume}(B(p.r)) \sim \frac{4\pi}{3}\left(r^3 - \frac{1}{30}R(p)r^5\right)$$

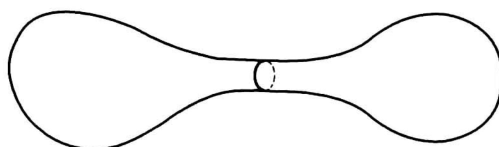
其中 $B(p,r)$ 是以 p 为圆心,以 r 为半径的球,$R(p)$ 是 p 点的数量曲率.





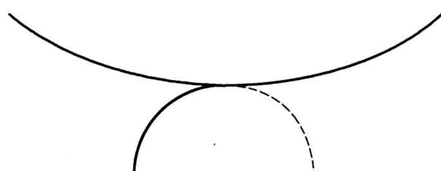
Ricci **曲率**

二维哑铃型曲面



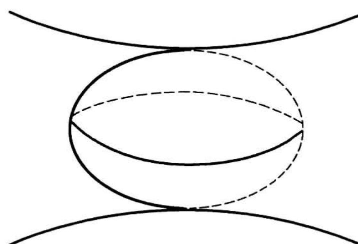
Ricci **曲率**

二维鞍马型负曲率曲面



Ricci 曲率

可是，在三维，一个颈有下面的形状：



这个片是具有很大正曲率的二维球面，所以颈部的数量曲率可以是正的.





爱因斯坦方程动力学

粗略地说,质量密度由空间的数量曲率加上动量密度组成.爱因斯坦动力方程迫使黑洞的形成,将空间分为两部分:数量曲率为正的部分空间和可能具有黑洞的部分空间.一般来说,在黑洞视界以内,拓扑趋向于容许负曲率结构.

重力理论中有两个量支配空间的动力学:度量与动量.动量很难控制,所以目前很难用广义相对论的爱因斯坦方程来研究空间的拓扑.

Hamilton 方程

1979 年,Hamilton 发展了新的方程来研究空间的变动.Hamilton 的方程是如下的

$$\frac{\partial g_{ij}}{\partial t} = -2R_{ij}$$

与重力驱动空间不同,他用 Ricci 曲率来驱动,这类似于热扩散.热传导方程具有使空间光滑的性质.它能够将热源瞬间传递到空间上的任何一点.

这个方程也被物理学家在同一时期考虑(首先出现在 Friedan 的论文里),不过观点有很大不同.

奇点

另一方面,整体拓扑与方程中由于曲率产生的非线性项确实将空间部分区域变为点,出现空间的拓扑塌陷.我们称这种点为空间的奇点.

1982 年时,Hamilton 在这个方程的研究方面发表了第一篇文章.从正曲率空间开始,他证明了,在他的方程支配下,在作保持体积不变的膨胀以后,空间不会遇到任何奇点,这就导致曲率在每个方向都是常数的空间.这种空间可以是三维球面,也可以是球面在有限等距群作用下的商.





看到 Hamilton 的定理后,我确信 Hamilton 的方程正是完成几何化纲领所需要的方程.

(在 Hamilton 的文章发表以后不久,出现了 Huisken 用平均曲率形变凸曲面的文章.平均曲率流方程是理解 Hamilton 方程的一个很好的模型)

我们建议用他的方程不断地改变三维空间,最后会将空间分解.这将会导致 Kneser,Jacob-Shalen, Johannson 的拓扑分解定理.我们希望哈密尔顿方程的渐近状态会分解成几个部分,或者塌陷,或者产生满足爱因斯坦方程的结构.

在三维空间中,爱因斯坦结构是常曲率的.可是,形变会产生奇点.主要的问题是找到描述所有奇点的办法.以下我们将介绍这个重要的发展.

Hamilton 纲领

Hamilton 的想法是通过拓扑手术把奇点除去,在手术以后继续他的方程.如果再次发展出奇点,则重复手术,继续前进.

如果我们可以证明在任意有限时间段内,只需做有限次手术,并且 Hamilton 方程的解的长时间行为得到了很好的了解,那么我们就能够识别出初始流形的拓扑结构.所以,Hamilton 的纲领如果能够成功实施,将会导致庞加莱猜想与 Thurston 猜想的证明.

Hamilton 的贡献的重要性与创造性永远不会被高估.这个领域里的任何专家都会认可 Hamilton 是整个理论最主要的贡献者.





Hamilton 纲领

2002 年 12 月,Perelman 说:

"遵循 Hamilton 纲领将会推出闭三维流形的几何化猜想."

"在这篇文章中,我们完成 Hamilton 纲领中的一些细节."

现在我们将根据年代发展,描述 Hamilton 的纲领.分成几个阶段:

I.先验估计

早在 20 世纪 90 年代,Hamilton 系统地发展理论,来理解奇点的结构.在我的建议下,他证明了当曲率为非负时,他的流的李伟光 - 丘成桐型估计(李伟光 - 丘成桐 - Hamilton 估计).这一估计提供了 Hamilton 方程行为的先验控制.

先验估计是证明非线性微分方程存在性定理的关键.一个直观的例子可以解释如下:一位导弹工程师设计导弹的轨线,他可能需要知道发射十分钟后导弹的位置与速度.由于风速的改变,他的估计可能很不相同.可是只要这个估计在一定的精度范围内,他就可以知道如何去设计导弹.如何决定这个精度范围,这就叫做先验估计.

李伟光－丘成桐－Hamilton 估计

在证明非线性微分方程的过程中,我们需要找到一些量的先验估计,来控制这个方程.在哈密尔顿流的情形,重要的量是数量曲率 R.数量曲率的相对大小对于奇点结构的研究非常关键.这是李伟光－丘成桐－Hamilton 估计给出的.

对任意 1－形式 V_a 我们有

$$\frac{\partial R}{\partial t}+\frac{R}{t}+\nabla 2_a R\cdot V_a+2R_{ab}V_aV_b\geqslant 0$$

特别的,$tR(x,t)$的每一点都是随时间递增的.





拼挤估计

在应用这种估计研究奇点结构的过程中，Hamilton 发现(lvey也独立得到)三维空间上他的方程的一个曲率拼挤估计.这使得他发现，奇点的一个邻域与非负曲率空间相似.

于是在附加了非塌陷条件下，Hamilton 得到了最可能出现的奇点的结构.只有一类奇点他无法确定其存在性(他将这类奇点称为雪茄型).

II. Hamilton 关于几何化的工作

1995 年，他发展了用常平均曲率曲面叶化的几何手术的方法来研究具有正迷向曲率四维流形的拓扑.

1996 年，Hamilton 继续分析在合适正则性条件(他称之为非奇异解)下的时空结构.特别的，他证明容许非奇异解的三维空间满足几何化猜想.

这些惊人的工作是建立在对几何与非线性微分方程的深刻分析基础上的.这两篇文章令人深信,几何化纲领可以用 Hamilton 的方法在可期待的一段时间内加以破解.

Hamilton 工作中的主要组成部分

在他深刻的分析中,需要如下几个主要部分:

1.他自己在郑绍远-李伟光-丘成桐 1981 年得到的单射半径估计工作基础上证明的关于空间收敛性的紧性定理.

2.Mostow 刚性定理的量化改进,刚性定理说具有有限体积的三维空间上至多只有一个常负曲率度量.





3.用环面分解空间的过程中,他需要证明环面是不可压缩的.他的证明依赖于 Meeks - 丘成桐和 Schoen - 丘成桐发展的极小曲面理论.Meeks - 丘成桐的工作得到了三维拓扑的 Smith 猜测证明中需要的一个等变 Dehn 引理.Schoen 丘成桐的工作与他们证明广义相对论中正质量猜测有关.

III. Perelman 的突破

在前面提到的这些工作出来以后,Hamilton 的方法给庞加莱猜想与几何化猜测的证明带来了一片光明.主要的困难在于如何利用局部曲率的界来对单射半径进行某种控制,以理解奇点的结构与奇点手术的过程.

2002 年 11 月,Perelman 推出预印本"The entropy formula for the Ricci flow and its geometric applications".在引言中,他说他的主要目的是执行 Hamilton 的纲领.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

与 1986 年李伟光 - 丘成桐的工作相平行的，Perelman 用路径积分引入了一个时空距离函数，用来验证一般的非塌陷条件. 令 σ 是任意链接 p 到 q 的时空道路，我们定义作用

$$\int_0^T \sqrt{s}(R + |\sigma(S)|^2)\mathrm{d}s$$

在所有 p 与 q 之间的道路取极小值，我们定义了时空距离函数.

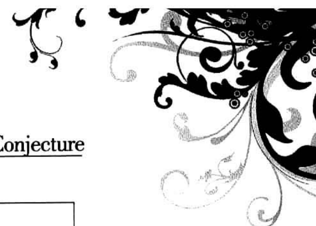
Perelman 定义他的约化体积

$$\int (4\pi\tau)^{-\frac{n}{2}}\exp\left\{-\frac{1}{2\sqrt{\tau}}L(q,\tau)\right\}$$

并注意到在 Hamilton 方程下，约化体积是递减的.

Perelman 说："李伟光 - 丘成桐对一个线性抛物方程赋以'长度'，与我们的情况几乎完全一样."

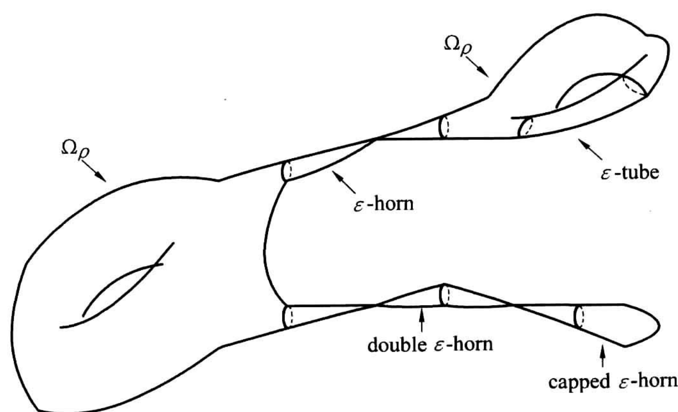




比例尺论证方法

Perlman 进一步发展了一个重要的,改进的比例尺论证方法,来完成 Hamilton 对于奇点的分类,得到了一致并整体的奇点结构定理.

奇点结构

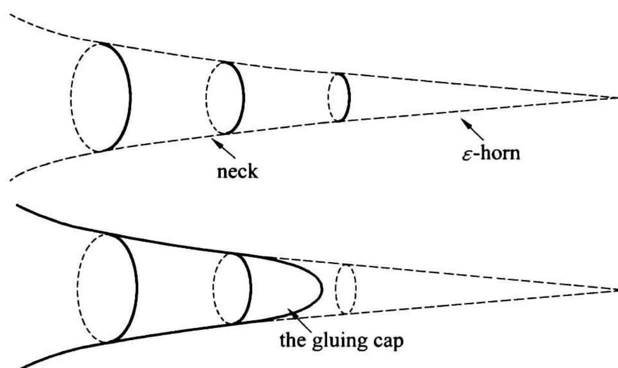


Hamilton 的几何手术

现在我们需要一个施行几何手术的办法.在 1995 年,Hamilton 已经对四维空间开创了一种手术过程,给出了施行这种手术的具体方法.

我们可以检验,Hamilton 的几何手术对三维空间也成立.手术的步骤介绍如下:

Hamilton 的几何手术





手术时间的离散性

真正的挑战在于证明每个有限时间段内只进行有限次手术. 问题在于,当我们在进行每次手术时,可能会引入误差,积累到一定程度手术就会越来越频繁.

2003 年 3 月, Perelman 推出了另一篇预印本,题为"Ricci flow with surgery on three manifolds",其中他改进了 Hamilton 的几何手术过程,随着时间演进,手术的精度也不断提高.

Perelman 引入了比例尺论证方法来研究手术时间离散性问题.

手术时间的离散性,比例尺论证

当对 Hamilton 方程的手术解采用比例尺论证,我们遇到如何应用 Hamilton 紧性定理的困难,因为这个紧性定理只对光滑解有效.

克服这个困难的想法由两部分组成:

1.(Perelman):选择截断半径充分小,将手术区域远推.

手术时间的离散性——比例尺论证方法

2.(曹怀东 - 朱熹平):建立了三个关于手术解的时间延展的结果,使得 Hamilton 的紧性定理仍然可用.为了达到这个目的,他们需要对手术区域的延伸有深刻的了解,这用到陈兵龙 - 朱熹平关于非紧流形上 Hamilton 方程解的唯一性定理.





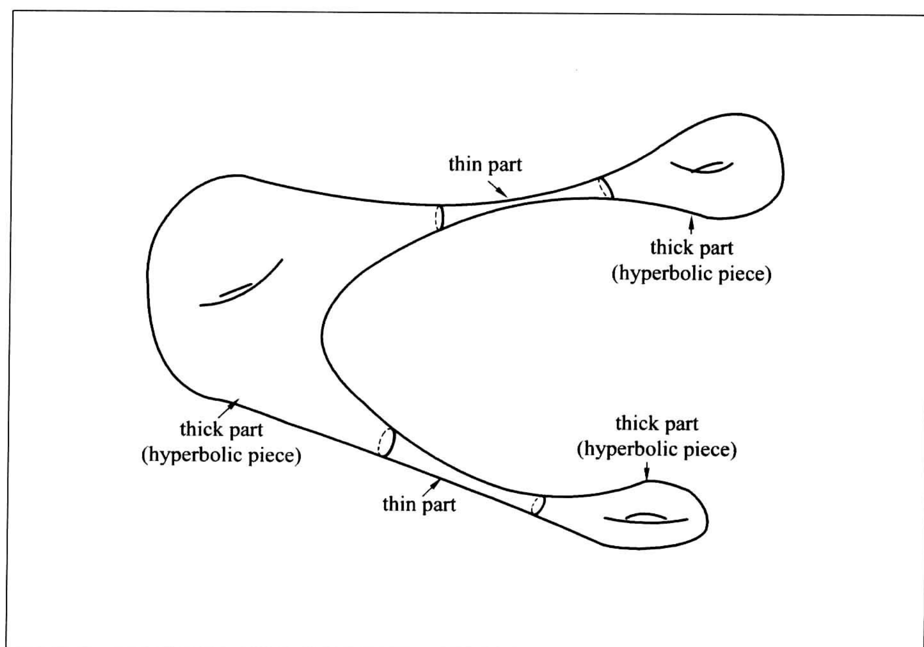
总结庞加莱猜想证明

一旦知道手术相对于时间是离散的,我们可以完成正数量曲率三维空间的分类,这是 Schoen - 丘成桐最早研究的问题.

更重要的,对单连通三维空间,结合 Colding-Minicozzi(2005)的有限时间消亡性结果,这就提供了庞加莱猜想的完整证明.

IV.几何化猜测的证明:粗细分解

为了研究一般空间的结构,我们仍然需要分析 Hamilton 方程的手术解的长时间行为.正如我们在 II 中所提到的,Hamilton 在 1996 年研究了一些特殊光滑解的行为.



粗细分解

Hamilton 证明任意三维非奇异解容许一种粗细分解,其中粗的部分由有限多个双曲空间组成,而细的部分塌陷.通过改进 Schoen - 丘成桐的极小曲面理论,Hamilton 进一步证明双曲片的边界是不可压缩环面.所以,任何非奇异解都是可以几何化的.



粗细分解

虽然非奇异的假设有局限性,但 Hamilton 的想法与论证在 Perelman 的工作中起了很关键的作用,特别用来分析一般手术解的长时间行为.

这样,Perelman 宣称可以用粗细分解来给出 Thurston 几何化猜测的证明.

粗细分解

对于粗的部分,基于李伟光 - 丘成桐 - Hamilton 的估计,Perelman 建立了一个关键的椭圆型估计,使得他可以证明粗的部分由双曲片组成.对细的部分,由于他只能得到截面曲率的一个(局部)下界,他宣布了一个新的塌陷结果.假设这个新的塌陷结果成立,Perelman 认为手术解与 Hamilton 工作中的非奇异解具有相同的长时间行为,这个结论可以推出 Thurston 几何化猜想的证明.

粗细分解

虽然 Perelman 期望的这个新的塌陷结果还未见诸文献，Shioya - Yamaguchi 在闭空间的特殊情形发表了一个塌陷结果的证明．最近，曹怀东 - 朱熹平在前面工作的基础上给出了 Thurston 几何化猜想的一个完全证明．

后记

在 Perelman 的工作中，许多关键的证明思想只是作了勾画或略述，而经常缺少完全的细节．最近曹怀东 - 朱熹平 2005 年提交给《亚洲数学杂志》(*Asian Journal of Mathematics*)的文章给出了庞加莱与 Thurston 几何化猜想的第一个完整与详细的描述．他们在自己工作基础上，给出了 Perelman 工作的几个步骤的新证明．





过去三年中,许多数学家试图探究:Hamilton 与 Perelman 的想法是否可以结合起来? Kleiner 与 Lott (2004)在网上公布了关于Perelman 部分工作的注记.不过这些注记与完整的证明相差甚远.在曹怀东-朱熹平的工作在 2006 年 4 月宣布以后,Kleiner 与 Lott在 5 月底公布了一篇比他们在 2004 年完整的笔记.他们的方法与曹怀东-朱熹平的方法有所不同.要完全理解他们的笔记还需要一些时间,因为在其中几个关键的地方仍然非常粗略.

Hamilton 纲领的成功是过去三十年中几何分析学家集体努力的成果.这应该被看做几何分析学科伟大的成就,它的奇妙之处在于只用几何与分析就能够证明极度困难的拓扑学定理.

Hamilton方程是一个复杂的非线性偏微分方程组.这是数学家们第一次能够理解复杂的偏微分方程组的奇点和演化.

类似的方程组在自然现象中比比皆是.解决Hamilton方程的方法将给这些自然的方程,如Navier-Stokes方程和爱因斯坦方程的研究带来曙光.

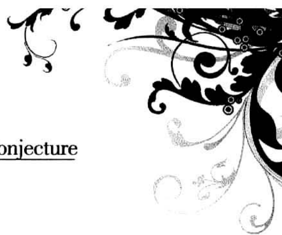
Hamilton方程的数值计算应当在计算机图形中有用的,在二维图形的研究中已由顾险峰-王雅琳-丘成桐所证实.

庞加莱(Poincaré):

"思想仅是漫漫长夜中的一个闪光,但这闪光意味着所有一切."

庞加莱在1904年的闪光已照明了20世纪拓扑学主要部分.庞加莱还开创了黎曼(Riemann)曲面的几何与分析理论.这个理论是20世纪所有数学发展的主要支柱之一.我相信三维空间的完全理解将在21世纪中起到类似的作用.





附录Ⅷ　几何分析

——丘成桐

这是我在中国数学会成立七十周年年会上的讲稿.我以此纪念我的老师陈省身先生,他在去年不幸去世了.

在我读研究生期间,我的研究选题十分广泛.我[538]首先研究的是非正曲率流形的基本群.从第二年起,我开始涉足流形上的微分方程.开始陈先生对我的这部分研究工作并未太过关注,几年后他方始觉察到这些工作的重要性.事实上,1972年陈先生在伯克利开设的课程上讲授了Calabi教授关于仿射几何的工作,我从郑绍远那里了解了这些鼓舞人心的课程.1973年,郑绍远和我开始研究陈先生课程中所提到的一些问题.我们没有注意到Pogorelov, Calabi和Nirenberg这几位大几何学家也在研究同样的问题.我们解决了Calabi关于虚(Improper)仿射球面的猜想,这着实令我们兴奋不已.可是很快我们就了解到Pogorelov [398]已在我们之前不久发表了文章.他用不同的方法证明了这个猜想.然而,我们的想法对于处理仿射几何中的其他问题也很有效.同时我对于Monge-Ampère方程的认识也在这些年中得到了很大的拓宽.陈先生对我的工作非常欣赏,特别是对我[543]在1976年完成的关于Kähler-Einstein度量的Calabi猜想的工作.当时我在斯坦福大学,陈先生建议伯克利数学系聘我为教授.我在1977年到伯克利访问一年,开设了几何分析方面的课程,着重介绍等距嵌入问题.

陈先生提名我在赫尔辛基国际数学家大会上作一小时报告.我的演讲[544]很成功.虽然陈先生对于我不愿留在伯克利有些不太高兴,但他还是建议我到普林斯顿高等研究所(IAS).其实在我接受IAS的职位前,1979年Borel邀请我在IAS组织了“几何年”的学术活动.这是令人振奋的一年,因为几何分析领域的大部分数学家都来了.

1979年,我接受华罗庚教授的邀请访问中国.我给了一系列关于Sacks-Uhlenbeck的冒泡过程(bubbling process) [413]的讲座.我建议中国数学工作者可以将类似的论证过程应用到界定两个具有相同常平均曲率曲面的Jordan曲线(作为公共边界).我觉得这是进入几何分析这个美妙领域的一个很好的习题.这个问题由王光寅教授[251]的一群学生作了研究.不幸的是,这也导致了中国数学会六十周年年会上的一场丑陋的争斗.由于丁姓学生的诬陷,王教授黯然

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

离开了数学界.这一事件极大减慢了过去十年中几何分析这一漂亮的学科在中国的发展.

1980年,陈先生决定在中国大范围发展几何分析.他倡议每年在中国举办微分几何与微分方程的国际会议.第一年,一大批杰出的数学家齐聚北京开设讲座(见[107]).我作了几何学中公开问题的演讲[546].可是中国数学家真正着手研究其中的一些问题,比我想象的要晚许多.

令陈先生失望的是,他在中国发展微分方程与微分几何的热情并未激发起他希望看到的欣欣向荣的景象.大多数国内数学家是学分析出身,可是几何功底很弱.几何分析的目标在于理解几何,但这一点并未受到足够的重视.另一方面,钟家庆的工作(见[562, 372, 373])是值得称道的.不幸的是,他在二十年前就去世了.微分几何的主要研究力量出自陈先生的学生与苏步青教授晚期的学生.谷超豪[208]教授做了出色的开创性工作,他在基于不同符号差(Signature)的度量上考虑调和映照.这种映照称为波映照.洪家兴(见[239, 220])在 $\mathbf{R}^3$ 中曲面的等距嵌入问题上作了一些有价值的工作.过去五年中,香港中文大学的研究中心在谭联辉和朱熹平的领导下,在 Hamilton 的 Ricci 流方面做出了第一流的工作(见[87,88,91,92,80]).

为了响应陈先生推动国内几何分析的号召,我向中国的数学同行们介绍一下我个人的一些观点.

微分几何的研究在我有幸参与的这段时间内一直发生着重大的进展.这篇文章的目的是与读者分享这些进展带来的喜悦.它不可能涉及这个学科的各个分支.我优先考虑的是与我的研究领域比较接近的工作.这样,我不参与其中的领域的工作就会少讲一点.尽管如此,我还是希望以一个过来人的亲身经历让读者,特别是那些年轻人了解几何分析这一领域的起源与三十年来的发展.我要感谢傅吉祥的帮助,特别是他对参考文献的检索.

我首先列出在如下各个不同时期对几何分析学作出主要贡献的数学家.

Ⅰ. 1972 ~ 1982 年: Atiyah, Bott, Singer, Calabi, Nirenberg, Pogorelov, Schoen, Simon, Uhlenbeck, Donaldson, Hamilton, Taubes, Stein, Fefferman, Siu(萧荫堂), Caffarelli, Kohn, S. Y. Cheng(郑绍远), Kuranishi, Cheeger, Gromoll, Harvey, Lawson, Gromov, Aubin, Patodi, Hitchin, Guillemin, Melrose, Colin de Verdière, M. Taylor, Bryant, W. Hu(伍鸿熙), R. Green, Peter Li(李伟光), Phong, Wolpert, Pitts, Trudinger, Hildebrandt, Kobayashi, Hardt, Spruck, Gerhardt, White, Gulliver, Warner, Kazdan.

Ⅱ. 1983 ~ 1992 年: 1983 年 Schoen 和我开始在 IAS 讲授几何分析.当时为我们作笔记的主要是钟家庆.这个课程 1985 年在 San Diego 分校继续进行.在 1985 到 1986 年期间,张恭庆和丁伟岳来美国为我们作笔记.《微分几何》这本专





著是 1989 年在中国出版的[434].这影响了整整一代中国数学家投身到几何分析这个学科中来.在此期间,我的学生中有许多人对这个学科作出了贡献.包括 Treibergs, Parker, Bartnik, Bando, Saper, Stern, Cao(曹怀东),Chow(周培能),Shi(施皖雄),Zheng(郑方阳)以及 Tian(田刚).

同时,Bismut, C. S. Lin(林长寿),N. Mok(莫毅明),J. Q. Zhong(钟家庆), Jost, Huisken, Jerison, Sarnak, Fukaya, Mabuchi, Ilmanen, Croke, Stroock, Price, F. H. Lin(林芳华),Zelditch, Christodoulou, Klainerman, Moncrief, Terng(滕楚莲),MichaelWolf, Anderson, LeBrun, Micallef, Moore, John Lee, A. Chang(张圣容),Korevaar 也在各自不同方向上作出了贡献.我们也要提到,在这个时期,前面一群数学家中也作出了许多重要的工作.例如,Donaldson, Taubes [476]以及 Uhlenbeck [509,510]关于一般流形上 Yang-Mills 理论的出色工作,使得 Donaldson [143]解决了四维拓扑中的著名难题.Donaldson [144], Uhlenbeck-Yau [512]证明了稳定丛上 Hermitian Yang-Mills 联络的存在性.Schoen [419]解决了 Yamabe 问题.

Ⅲ.1993 ~ 现在:有许多数学家加入到几何分析中来.包括 Kronheimer, Mrowka, Demailly, Colding, Minicozzi, T. Tao(陶哲轩),Thomas, Zworski, Eliasberg, Toth, Andrews, L. F. Tam(谭联辉),C. Leung, Y. Ruan(阮勇斌),W. Ruan(阮卫东),R. Wentworth, Grigorian, Saloff-Coste, J. X. Hong(洪家兴),X. P. Zhu(朱熹平),M. T. Wang(王慕道), A. K. Liu(刘艾克), K. F. Liu(刘克峰),X. F. Sun(孙晓峰),T. J. Li(李天军),X. J. Wang(汪徐家),Loftin, Y. Li(李岩岩),H. Bray, J. P. Wang(王嘉平),L. Ni(倪磊),P. F. Guan(管鹏飞). N. Kapouleas, Ozsváth, Szabó, Y. Li(李莹英).这一时期最重要的突破当数 Hamilton [218]在 1995 年关于 Ricci 流的工作.虽然我早在 1982 年就建议他用 Ricci 流来解决 Thurston 猜测.可是直到 Hamilton 这篇文章出来以后,我们才真正认识到用几何分析方法解决几何化纲领是完全可行的.(关键的一步是李伟光和我[309]首先研究的关于抛物方程的估计被 Hamilton 用到了 Ricci 流上[216, 217])2002 年,Perelman [389, 390]引入了全新的思想来解决几何化纲领中留下的关键步骤.希望陈兵龙 - 朱熹平[91,92]和曹怀东 - 朱熹平的工作能够最终完成整个纲领.在另一个方向上,我们看到了 Seiberg-Witten 理论的重要发展[533]. Taubes [482, 483, 484, 485]用他的不变量证明了伪全纯曲线计数的重要定理. Kronheimer-Mrowka [279] 证明了 Thom 猜测,即全纯曲线提供了代数曲面中表示同调的最低亏格的曲面(Ozsváth-Szabó 有一个对应的辛几何版本[387]).

既然整个几何分析学科的发展很强地依赖于过去,就让我们从历史的发展开始.下面是 1970 年以前的工作回顾,这些工作提供了许多富有成效的思想和

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

方法.

• 费马(Fermat)的变分原理(各种条件下的最短路径问题).

• 牛顿(Newton)与莱布尼兹(Leibniz)的微积分:物体的运动路径由自然界的规律支配.

• 欧拉(Euler)、拉格朗日(Lagrange):变分原理基础和偏微分方程的研究;流体方程与极小曲面方程的推导.

• 傅里埃(Fourier)、希尔伯特(Hilbert):函数关于特征函数的分解,谱分析.

• 高斯(Gauss)、黎曼(Riemann):内蕴几何的概念.

• 黎曼、狄利克雷(Dirichlet)、希尔伯特:用变分法解决调和函数的狄利克雷边值问题.

• 麦克斯韦(Maxwell):电磁场、规范场与作用力的统一.

• 克里斯托弗(Christoffel),勒维-齐维塔(Levi-Civita)、比安奇(Bianchi)、瑞奇(Ricci)流形上的微积分.

• 黎曼、庞加莱(Poincaré)、科伯(Koebe):黎曼面单值化理论,共形形变.

• 嘉当(Cartan):外微分方程组,纤维丛上的联络.

• 爱因斯坦(Einstein)、希尔伯特:爱因斯坦方程与希尔伯特作用.

• 狄拉克(Dirac):旋量(spinors),狄拉克方程,量子场论.

• 凯勒(Kähler)、霍奇(Hodge):凯勒度量与霍奇理论.

• 希尔伯特、科恩-弗森(Cohn-Vossen)、列伟(Lewy)、外尔(Weyl)、霍普夫(Hopf)、波戈里洛夫(Pogorelov)、埃非莫夫(Effimov)、尼伦伯格(Nirenberg):基于分析的三维空间中的整体曲面理论.

• 魏尔斯特拉斯(Weierstrass)、黎曼(Riemann)、勒贝格(Lebesgue)、库朗(Courant)、道格拉斯(Douglas)、拉多(Radó)、摩里(Morrey):极小曲面理论.

• 高斯、格林(Greeen)、庞卡莱、肖德尔(Schauder)、摩里:位势论,椭圆方程正则性理论.

• 外尔、霍奇、小平邦彦(Kodaira)、德拉姆(de Rham)、伯格曼(Bergman)、米尔格兰姆-罗森布鲁姆(Milgram-Rosenbloom)、阿蒂亚-辛格(Atiyah-Siniger):德拉姆-霍奇理论,积分算子,核函数,指标理论.

• 庞特里亚金(Pontrjagin)、陈省身、艾伦多弗-韦伊(Allendoerfer-Weil):曲率形式定义的整体拓扑不变量.

• 博赫纳-小平(Bochner-Kodaira):基于曲率考虑的上同调群的消没.

• 德乔治-纳什-莫泽(De Giorgi-Nash-Moser):高维椭圆方程的正则性理论与散度型抛物方程.

• 小平、摩里、荷曼德尔(Hörmander)、科恩(Kohn)、安德雷奥蒂-维森蒂尼(Andreotti-Vesentini):复流形的嵌入,$\bar{\partial}$-诺伊曼问题.





• 小平－斯潘塞(Spencer)、纽兰德(Newlander)－尼伦伯格:几何结构的形变.

• 费德勒－弗莱明(Federer-Fleming)、阿姆格伦(Almgren)、阿勒德(Allard):高维多形体(varifold)与极小簇.

• 伊尔斯－桑普森(Eells-Sampson):到非正曲率流形调和映照的存在性.

• 卡拉比(Calabi):仿射几何与关于凯勒－爱因斯坦度量的猜想.

我现在将要解释 20 世纪 70 年代以来取得的成果,我将从对于几何与分析相互关系的系统研究开始.整个内容依赖于我个人的品位,可能遗漏了其他许多学者的贡献.

1 几何中函数的构造

我们有如下的基本原则[544]:

发展线性或非线性分析的目的是为了更好地了解空间的几何或组合结构.在这个过程中,几何学为分析学提供了深刻的洞察力.一个重要的方针就是,由空间结构定义的特殊函数空间可以反过来用来定义空间自身的结构.

代数几何学家用有理函数环定义了代数簇的 Zariski 拓扑.

在微分几何中,我们应该能够通过定义在其上的函数,把流形的度量与拓扑信息提取出来.自然地,这些函数应该由几何构造或由几何的基本结构给定的微分方程来加以定义.(积分方程没有被广泛使用,因为将局部几何与整体几何联系起来的想法与微分方程有更大的关系)函数的自然推广包括微分形式、旋量(spinors)与向量丛的截面.

微分形式或向量丛的概念的对偶是子流形或叶化(foliation).从支配它们的微分方程(大多是来自变分原理)来看,我们有极小子流形或全纯闭链.自然地,它们的性质或所有这些对象的全体支配着原本流形的几何.一个非常好的例子是流形的环路(Loop)空间上的 Morse 理论.

我们现在讨论构造几何学中感兴趣的函数与张量的各种方法.

1.1 来自于外围空间的多项式

如果流形等距嵌入到欧氏空间,欧氏空间中多项式函数在其上的限制就是一类自然定义的函数.可是,等距嵌入通常不是刚性的,所以如此构造出来的函数一般不太有用.

另一方面,如果一个流形以典则的方式嵌入到欧氏空间中并且这个子流形的几何由欧氏空间的某个线性变换群所定义,那么限制在子流形上的多项式确实起到了很重要的作用.郑绍远和我(1974, 1975 年)[102, 106]发展了这些函

数的几个重要的梯度估计来控制这些子流形的几何.

第一个重要的定理是关于 Minkowski 空间中的类空超曲面 M.一个非常重要的问题是:因为 $\mathbf{R}^{n,1}$ 上的度量是 $\sum(\mathrm{d}x^i)^2-\mathrm{d}t^2$,它到 M 上的限制不一定完备,即使当 M 在欧氏度量下完备.为了证明在这两个度量下平均曲率可以控制的超曲面上这两个完备性概念的等价性,郑绍远和我证明了限制在超曲面上的函数

$$\langle X,X\rangle=\sum_i(x^i)2-t^2$$

的梯度估计。

通过坐标系的选取,可以假定函数 $\langle X, X\rangle$ 在 M 上是正的和逆紧的.对 M 上的任意正值逆紧函数 f,如果我们证明如下的梯度估计

$$\frac{|\nabla f|}{f}\leqslant C$$

其中 C 与 f 无关,那么我们就可以证明 M 上的度量是完备的.这可以通过对上述不等式积分得到

$$|\log f(x)-\log f(y)|\leqslant Cd(x,y)$$

所以当 $f(y)\to\infty$, $d(x,y)\to\infty$.一旦我们知道度量是完备的,我们就证明了 Bernstein 定理,即极大类空超曲面必是线性的.这项工作后来被 Treibergs [506],C. Gerhardt [181] 和 R. Bartnik [28] 推广到更一般时空中的超曲面.(如果我们假设爱因斯坦方程的初始值是非奇异的,理解在一般时空中有爱因斯坦方程演化的极大类空超曲面的叶化的性质仍然是一个重要的问题)

另一个重要的例子是对仿射空间 A^{n+1} 中的仿射超曲面 M^n 的研究.这些包括虚(improper) 仿射球面

$$\det(u_{ij})=1$$

其中 u 是一个凸函数与双曲仿射球面

$$\det(u_{ij})=(-\frac{1}{u})^{n+2}$$

其中 u 是凸区域 Ω 的凸函数且限制在边界 $\partial\Omega$ 上等于零.注意这些方程描述了仿射法线平行或收敛到一个点的超曲面.

对仿射几何来说,有一个定义在 M 上的仿射不变度量,即

$$(\det h_{ij})^{-\frac{1}{n+2}}\sum h_{ij}dx^i\wedge dx^j$$

其中 h_{ij} 是 M 的第二基本形式.一个基本的问题是,这个度量是否完备.

通过选择 A^{n+1} 中的一个坐标系,使得高度函数成为一个定义在 M 上的逆紧正函数.高度函数的梯度估计给出了证明仿射度量完备性的一个途径.郑绍远和我[106] 的确找到了与上面相类似的估计.

一旦知道了仿射度量的完备性,仿射球面的性质证明起来就变得极为容





易,包括一些卡拉比的猜测.例如,我们证明了虚仿射球面是抛物面,并且每个真凸锥容许一个双曲仿射球面的叶状结构.这个命题的前半部分最早是由Jörgens [252], Calabi[67]和Pogorelov[398]证明的.反之,我们也证明了每个双曲仿射球面渐近于一个凸锥.(直到十年以后国内某数学家重复郑绍远和我的估计,宣布得到了同样的结果)就在不久前,Trudinger与汪徐家[508]解决了关于仿射极小曲面的Bernstein型问题,从而证明了陈省身的一个猜想.他们还找到了维数大于等于10的反例.这些结果也使我们更清楚地看到,可以通过研究四阶椭圆方程来更好地理解仿射极小曲面.

这种利用某些自然定义的函数的梯度估计的论证方法,也被我用来证明郑绍远与我所构造的有界拟凸域上的凯勒-爱因斯坦度量的完备性[104].(在我与莫毅明的文章[371]中证明了这个定理的逆,即如果凯勒-爱因斯坦度量是完备的,那么这个域是拟凸的)郑绍远和我的证明方法是基于极大值原理.在这个方法发展之前,利用De Giorgi-Nash的想法Bombieri-Giusti [45]作出了关于极小图梯度估计的深刻工作.后来,N. Korevaar [274]找到了一个基于极大值原理的Bombieri-Giusti工作的一个简单的证明方法.

基本原则:要控制一个几何度量,可以通过找到一个我们所熟知的函数,给出这个函数的梯度或更高阶的估计.

在1982年,郑绍远、李伟光和我[100]用这个原理给出了欧氏空间、球面和双曲空间中极小子流形的一个热核比较定理.因为任意$\mathbf{C}P^n$中的复子流形都可以提升到S^{n+1}中的一个极小子流形,所以很多年之后李伟光-田刚[304]关于$\mathbf{C}P^n$复子流形的类似工作只不过是一个推论.

线性函数的另一个非常重要的性质是说,当它限制到球面中的一个极小超曲面上时,它自动就成为一个特征函数.当超曲面是嵌入的,我猜测第一特征函数是线性的,相应的第一特征值等于n(见[546]).虽然这个猜测还没有完全得到证明,但Choi-王蔼农[112]的工作在很大程度上支持了这个猜测的正确性.他们证明了第一特征值有一个只依赖于n的下界.这个结果已足够被Choi-Schoen[110]用来证明S^3中的嵌入极小曲面的一个紧性结果.

从外围空间还可以构造另一类函数.这些就是定义在球面上的超曲面的支撑函数.著名的Minkowski问题约化为解关于支撑函数的Monge-Ampère方程.这个问题被Nirenberg [382]、Pogorelov [397]、Cheng-Yau [103]所解决.关于给定主曲率的对称函数的相应问题已经由许多人做了研究,包括Pogorelov [399], Caffarelli-Nirenberg-Spruck [65],管鹏飞(见[210, 209]),关波,Gerhardt [182]等.曲面到三维空间的等距嵌入问题可以写成支撑函数的Darboux方程.对局部嵌入问题,有林长寿[320, 321]的深刻工作,以及最近Han-Hong-Lin [220]的工作.负曲率曲面的整体嵌入问题由洪家兴作了研究[239].

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

空间形式的子流形如果它的法丛是平坦的，且主曲率沿着平行法向量场是常数，则称为是等参的. 这类子流形由 E. Cartan 开始研究. 最近有滕楚莲和 Thorbergsson 的工作(见 Terng 的综述文章[488] 和 Thorbergsson [493]). 滕楚莲[487] 还把欧氏空间中的嵌入连通双曲空间和孤立子理论联系起来. 与几何相关的 Lax 对和环路群的漂亮理论也得到了发展.

1.2 函数的几何构造

当流形不能嵌入到线性空间中，也有办法构造与度量结构相适应的函数. 显然，距离函数是第一个可用的主要函数. 从距离函数，我们可以构造 Busemann 函数如下：

给定一条测地射线 $\gamma:[0,\infty)\to M$ 使得

$$d(\gamma(t_1),\gamma(t_2)) = t_2 - t_1$$

其中 $\|\frac{\mathrm{d}\gamma}{\mathrm{d}t}\|$，我们定义

$$B_\gamma(x) = \lim_{t\to\infty}(d(x,\gamma(t)) - t)$$

这个函数推广了线性函数的概念. 对双曲空间形式，它的水平集合定义了极限球面 (horosphere). 对正曲率流形，它是凹的. Cohn-Vossen(曲面情形) 和 Gromoll-Meyer [196] 用它证明了正曲率完备非紧流形微分同胚于 $\mathbf{R}^n$.

Busemann 函数的一个重要性质是它在完备非负 Ricci 曲率流形上在分布意义下是上调和的. 这是证明 Cheeger-Gromoll 分裂原理的关键. 各种不同形式的分裂原理在流形结构研究中起了重要的作用. 当我[543] 证明卡拉比猜测时，我和其他一些学者用分裂原理证明了具有非负第一陈类的凯勒流形的结构定理. (结构定理的证明方法是 Calabi [66] 给出的，要用到第一 Betti 数的性质. Kobayashi [267] 和 Michesohn[363] 写出了严谨的步骤，Beauville [33] 有一篇综述介绍这方面的发展)

1974 年，我用 Busemann 函数给出了非负 Ricci 曲率完备流形的体积估计[541]. 这项工作被 Gromov [201] 推广到测地球的体积估计. 这些工作在最近 Perelman[389, 390] 有关 Hamilton 的 Ricci 流的工作中有重要应用.

如果我们考虑 $\inf_\gamma B_\gamma$，其中 γ 跑遍流形上从一点出发的所有测地射线，我们有可能得到流形的一个逆紧的穷竭函数. 当 M 是有限体积的完备流形，并且曲率介于两个负常数之间，萧荫堂和我[459] 证明了这个函数确实给出了流形的一个凹的穷竭函数. 如果流形是凯勒的，我们能够证明，可以通过在每个端添加一个点来得到一个紧致的复簇.

我也断言每个具有有界曲率，有限体积和有限拓扑的完备凯勒流形可以紧化为一个紧致复簇. 我在 1982 年建议莫毅明与钟家庆研究这个问题，他们在这





个方向上得到了重要的结果[373].(莫毅明 - 钟家庆给出的紧化不是典则的,我们希望找到 Bore-Baily 紧化[20] 的一个代数几何的类似结论,这样我们就可以用紧化空间的相交上同调来研究 L^2 上同调) 回忆一下把 L^2 上同调和 Borel-Baily 的相交上同调统一起来的 Zucker 猜想,这已经被 Saper-Stern [415] 和 Looijenga [337] 证明.如果至少对各种不同的模空间,存在类似的紧化,那么将是很有意义的.

除了到一个点的距离函数,我们也可以考虑到一个子流形的距离函数.在这个情形下,它的 Laplacian 将会涉及外围空间的 Ricci 曲率与子流形的平均曲率.这种函数可以用来作为构造极小子流形的闸函数(barrier). Schoen 和我[431] 用这个闸函数再构造极小曲面证明了任意完备正 Ricci 曲率三维流形微分同胚于欧氏空间.

如果我们考虑流形上的环路空间,我们可以取每个环路的长度,从而定义环路空间上的一个自然的函数,Morse 理论在这个函数的研究中有丰富的应用.自从 Poincare, Birkhoff, Morse, Ljusternik-Shnirel'man 等人的工作以来,利用环路空间上的 Morse 理论证明闭测地线的存在性有许多的研究.Klingenberg 和他的学生发展了十分有用的工具(见[266]).Gromoll-Meyer[197] 证明了无穷多条闭测地线的存在性.后来还有 Ballmann, Ziller, Hingston 和 Kramer 的工作(见[24, 227, 278]).当度量是 Finsler 的,最近 Victor Bangert 和龙以明[27] 证明了二维球面上两条闭测地线的存在性.(Katok [261] 给出的例子表明两条测地线是最优的) 在陈先生的领导下,鲍大维、沈忠民、莫小欢与吉敏等发展了 Finsler 几何(见[29]).

当流形具有负曲率时,曲线的长度函数与如下定义的位移函数有关:

如果 γ 是作用在一个非正曲率完备流形的万有覆盖基本群中的元素,我们考虑函数 $d(x,\gamma(x))$:对这个函数的研究可以给出紧致非正曲率流形的性质.例如,在我的博士论文中,我把 Preissman 定理推广为基本群的每个可解子群必是一个阿贝尔群的有限扩张,其中这个阿贝尔群是一个全测地平坦子环的基本群[538]. Gromoll-Wolf[198] 以及 Lawson-Yau[284] 证明了,如果这种流形的基本群没有中心,并且分裂为乘积,那么这个流形分裂为度量乘积流形.

当流形的曲率有界,Margulis 研究了使得 $d(x,\gamma(x))$ 很小的那些点,并证明了著名的 Margulis 引理.Gromov [199] 用这个引理大量研究了非正曲率流形的结构. 他 [200] 也对距离函数发展了 Morse 理论的思想(K.Grove 和 K.Shiohama 有一个初步的理论[207]),来比较一个测地球与一个更大测地球的相对拓扑,从而得到非负截曲率紧流形的 Betti 数的上界.(也可以允许负曲率,只要曲率有下界,可是估计值就要用到直径和曲率的下界)

Busemann 函数给出了检测流形在无穷远处的"角结构" 的方法.它可以用

来构造双曲空间形式的 Poisson 核.对具有有界和强负曲率的单连通完备流形，它可以用作解有界调和函数狄利克雷问题的闸函数，在无穷远处作卷积光滑化.这是 Sullivan[471] 和 Anderson [4] 的工作.对有界调和函数，Schoen 和 Anderson [5] 有深刻的 Harnack 不等式，并且识别出这种流形的 Martin 边界.另外还有 W. Ballmann [21] 的关于非正曲率流形 Dirichlet 问题的重要工作.

Martin 边界在各种对称空间紧化的情形由季理真和 Macpherson [250] 作了研究.对于负曲率流形的乘积，有 A. Freire [170] 关于 Martin 边界的漂亮工作.对秩一非正曲率完备流形，有 Ballman-Ledrappier [22] 和 Cao-Fan-Ledrapier [75] 的工作.这些工作与更高秩对称空间上的华罗庚算子有关.华发现了满足来自于华算子的额外方程的有界调和函数(见[242]).

1.3 由微分方程定义的函数

流形上最重要的微分算子是拉普拉斯算子.它的谱分解给出特征函数.所以调和函数是在几何中扮演重要角色的最简单的函数.

如果流形是紧致的，由极大原理知道调和函数必是常数.可是，当我们想要了解紧流形的奇点时，我们可以通过缩放和爆破的过程来得到非紧流形.那么调和函数就又可以起重要作用了.

完备流形上的第一个重要问题是 Liouville 定理.我在分析上最初的工作就是为了理解刘维尔定理的合理陈述.在 1971 年，我意识到可以很自然地证明非负 Ricci 曲率的完备流形上不存在非平凡调和函数[539].我也注意到反过来，如果完备流形具有强负曲率，并且是单连通的，那么我们可以解决有界调和函数的狄利克雷问题.

我推导正调和函数的梯度估计[539] 源于对复分析中 Schwarz 引理的合理解释. 事实上，在我得到调和函数的梯度估计之前，我已经推广了 Ahlfors-Schwarz 引理.广义 Schwarz 引理[545] 说，从一个 Ricci 曲率有下界的完备凯勒流形到全纯双截曲率以负常数为上界的 Hermite 流形的全纯映照，是距离减少的，其中常数只依赖于曲率的上界.这个推广被 Liu-Sun-Yau [333, 334] 用来在模空间几何学中找到了许多应用.

我所发现的梯度估计可以推广到特征函数.李伟光[301] 是最早用它来找到正 Ricci 曲率流形特征值估计的人.(如果 Ricci 曲率有一个正的下界，那么最早的估计是 Lichnerowicz 得到的)Li-Yau[307] 接着解决了用直径和 Ricci 曲率下界给出流形特征值估计的问题.Li-Yau 猜测了下界估计的精确值，被钟家庆 - 杨洪苍[562] 通过改进 Li-Yau 的方法所证明.陈木法 - 王凤雨[94] 发展了





概率论的方法来推导这些不等式.特征值的精确上界首先由郑绍远[97]得到.郑绍远的定理给出了一个很好的通过对函数的分析来提供几何信息的范例.作为他的定理的推论,他证明了如果一个紧致流形 M^n 的 Ricci 曲率大于等于 $n-1$,并且其直径等于 π,那么流形等距于球面.他用到了 Lichnerowicz 和 Obata 的特征值下界估计.Colding[123]利用与第一特征函数性质相近的函数证明了一个拼挤定理:当 Ricci 曲率下界为 $n-1$,且体积与单位球足够接近,那么流形就微分同胚于球面.Colding-Cheeger[83, 84, 85]和 Perelman(见[61])等的详尽工作增进了我们对于流形 Hausdorff 收敛的 Gromov 理论的了解.他们使用的工具包括比较定理,Cheeger-Gromoll 分裂定理和 Colding 早先引入的思想.

经典的刘维尔定理有一个自然的推广:多项式增长的调和函数其实是多项式.受到这个事实以及多复变理论的启发,我提出一个问题,是否具有固定增长速率的多项式增长调和函数空间是有限维的,并且维数的上界只依赖于增长率[548].这个事实已经由 Colding-Minicozzi [124] 所证明,并被李伟光[302]加以推广.(函数可以替换为丛的截面)在一系列漂亮的文章中(见[305,306]),李伟光和王嘉平研究了与流形几何有关的调和函数空间.当调和函数是全纯的,他们构成一个环.我对这个环的结构很好奇,什么时候它是有限生成的?

这些工作大多可以推广到 Sobolev 不等式和 Poincaré 不等式成立的流形上去.(应该指出 Aubin[17, 19]和 Talenti[473]发现了欧式空间上各种 Sobolev 不等式的最佳常数)这些不等式都与等周不等式有关.这个方向上的第一项重要工作是 Cheeger [81] 作出的,他指出了等周不等式在特征值估计中的重要性.C. Croke [134] 在我[540]有关 Poincaré 不等式的工作基础上证明,Sobolev 不等式只依赖于体积,直径和 Ricci 曲率的下界.Cheng-Li-Yau [99] 应用 John Nash 的论证方法,给出了热核及其高阶导数的估计.在这篇文章中,我们得到了内射半径的一个估计,这个估计后来在 Hamilton 的 Ricci 流理论中发挥了作用.D. Stroock(见[380])应用 Malliavin 的随机变分理论给出了在一对点处(其中一个点是另一个点的割点)的一个重要的热核估计.

热核估计后来被 Saloff-Coste 以及 Grigor'yan 等人推广到具有多项式体积增长和体积增倍性质的完备流形上.因为它们都是拟等距不变量,所以这些分析可以应用到图或者离散群上.参看 Grigor'yan [193] 和 Saloff-Coste [417] 的综述文章.

另一方面,我所推导的原始的梯度估计是一个点点成立的不等式,在非线性理论中尤其有用.李伟光和我[309]在 1984 年得到了一个抛物情形的梯度估计.我们注意到它在热方程以及与时空中道路变分原理相关问题中的重要性.这些想法在 Hamilton [216, 217] 关于 Ricci 流的分析中扮演了重要的角色.

等周不等式是一个漂亮的课题.它有着悠久的历史.除了与特征值的关系,

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

它也揭示了深刻的流形结构信息.找到等周不等式的最佳常数是重要的问题.Pólya-Szegö[400], G. Faber(1923), E. Krahn(1925), P. Lévy(1951) 等作出了贡献.Gromov 推广了 Lévy 的想法,得到 Cheeger 常数的较好估计(见[203]).C. Croke [135] 以及曹建国与 Escobar [74] 在非正曲率的单连通流形的区域上进行研究.等周不等式对欧氏空间中的任意极小簇都成立,但最佳常数并不知道.Li, Schoen 和我[303] 在曲面的情形作了一些研究.E. Lutwa, Deane Yang 与张高勇在仿射几何的情形做了一些漂亮的工作(见[343, 344]).

给出拉普拉斯算子特征值的一个非常精确的估计在数学的许多领域都很重要.例如,我分别与 Paul Yang [537] 以及李伟光[308] 合作把 Szegö[472] 和 Hersch [226] 关于第一特征值上界的想法推广到更高亏格.我[551] 应用它证明,由算术群定义的黎曼面当它分支覆盖在球面上时,必定有一个相对高的次数.在一篇漂亮的文章中,N. Korevaar [275] 给出了黎曼面的第 n 个特征值 λ_n 的只依赖于亏格和 n 的上界.他的结果解决了我提出的一个猜测(见[546]).Grigor'yan, Netrusov 和我[194] 简化了他的证明,并给出了在极小曲面指标估界中的应用.此外还有 P.Sarnak(见[416, 249]) 关于理解这些黎曼面的特征函数的工作.他指出了其与数论的联系,并指出算术曲面上第 n 个特征函数的极大范数的估计在数论中是很有意义的工作.

特征函数有许多重要的性质在20世纪70年代作了研究.例如,郑绍远[98] 基于亏格给出了特征值重数的漂亮估计.这些想法被 Colin de Verdière [132] 用来把满足一些很好的组合性质的图嵌入到 $\mathbf{R}^3$ 中.结点域的连通性和拓扑是很有意思的问题.Melas [359] 对于凸平面区域,证明了第二特征函数的结点线一定与边界正好交于两点.目前关于结点域的数目所知甚少,仅有 Courant 的著名定理,即第 m 个特征函数的结点域的数目小于 m.

几何分析中有许多其他的想法可以被离散化并且应用到图论中.特别是图的谱理论.这些方向还有很多的问题可以做.金芳蓉,Grigor'yan 和我已经做了一些工作(见金芳蓉的综述文章[121] 中的参考文献).可是[121] 中的结果还远没有完善.另一方面,Margulis[347], Lubotzky-Phillips-Sarnak [340] 应用离散群和数论来构造扩张图.构造与分类这些扩张图的方法在计算机科学中都很重要.研究用谱方法给出任意图的一个好的分解也很重要.

有几个重要的问题与结点集和特征函数临界点数目有关.我曾作过一个关于结点集面积的猜想(见[546]),Donnelly-Fefferman [153],董瑞涛[152] 和林芳华[323] 在这个问题上做出了重要的贡献.特征函数的临界点数目很难计算.我[552] 成功证明了结点集附近临界点的存在性.Nadirashvili 证明了关于特征值及其特征函数的几个漂亮结论.他与合作者[238] 是最早证明维流形上的特征函数解的临界点集具有有限 H^{n-2}-Hausdorff 测度.后来,Han-Hardt-Lin [219] 给





了一个显式的估计.

如果有一个位势 V,那么 $-\Delta+V$ 的特征值也很重要.我与 Singer, Wong 和 Stephen Yau [450] 研究了间隙 $\lambda_2-\lambda_1$.我[553] 研究了间隙如何依赖于 $-\log\psi$ 的 Hessian 的低阶特征值,其中 ψ 是基态.当 V 是数量曲率,我和 Schoen 作了详细的研究.这些问题与共形形变,极小子流形稳定性等有自然的联系.(D. Fischer-Colbrie 和 Schoen [163], Micallef [361], Schoen-Yau [425, 431] 的关于稳定极小曲面的工作都依赖于对这个算子谱的理解) 抛物的情形出现在最近 Perelman 的工作中.

我们有 Fefferman, Phong, E. Lieb, Duistermaat, Guillemin, Melrose, Colin de Verdier, Taylor, Toth, Zelditch, Sarnak 从半经典分析的观点理解 Laplacian 谱的重要工作(见[161, 155, 225, 418]),他们的部分想法可以在 J. Keller 关于几何光学的分析中找到.Duistermaat 和 Hörmander [154] 关于奇点传播的基本工作也被广泛使用.对于决定什么条件下"可以听到鼓的形状" 这个困难的问题已经有了很多进展,比如 Melrose(见[360]),Guillemen [211], Zelditch [559] 等人的工作.(二维的反例由 Gordon-Webb-Wolpert [189], Wilson, Szaboó 给出)

调和函数或全纯函数的自然推广是具有联络的丛的调和或全纯截面.最重要的丛是余切丛的外形式幂.利用 Levi-Civita 联络,调和截面就是调和形式,由 DeRham 与 Hodge 理论,这些给出了上同调类的典则表示.调和形式的研究主要用到 Bochner 消没定理[43].可是我们目前的了解除了一维以外还是很贫乏.如果流形是凯勒的,微分形式可以进一步分解为(p,q) - 形式.Kodaira 消没定理[268] 给出了更强的信息.

为了实践通过函数论来决定流形结构的想法,在我还是研究生的时候,我就想要把黎曼面单值化定理推广到高维.那段时期,我在构思猜想时,受到 Greene-Wu[192] 中部分工作的影响.Greene 和伍鸿熙对流形是否为 Stein 的更感兴趣.

当完备凯勒流形是紧致的,并且具有正的双截曲率,这就是 Mori[374],以及 Siu-Yau [458] 所独立证明的 Frankel 猜想.论证都依赖于低次有理曲线的构造.当我们假设流形有非负双截曲率和正的 Ricci 曲率,那么,Mok-Zhong [372] 和 Mok[368] 证明流形是 Hermitian 对称的,其中用到了 Bando [25] 在博士论文中应用 Hamilton 的 Ricci 流到这种流形的低维情形的想法.

当完备凯勒流形是非紧的,且具有正的双截曲率,我猜测它必双全纯等价于 $\mathbf{C}^n$(见[546]).Siu-Yau [457] 用 Hörmander [241] 构造慢增长全纯函数的 L^2 方法在这个猜想的证明上作了第一个尝试.(注意 Hörmander 的方法可以追溯到 Kodaira 以及 Calabi-Vesentini [68] 的推广) 这篇文章中用到了奇异权函数,更加细致的论证由 Nadel [378] 和萧荫堂[456] 加以发展,被称为乘数层

(multiplier sheaf) 方法,在代数几何中有重要应用.

接着 Siu-Yau 的工作,Siu-Mok-Yau [369] 和 Mok [366, 367] 在假设曲率衰减和体积增长条件下做了一些工作.施皖雄[439, 440, 441] 开始引入 Hamilton 的 Ricci 流来研究我的猜测,并且作了基本重要的工作.后来还有曹怀东[72, 73], Chen-Zhu[89, 90] 和 Chen-Tang-Zhu [87] 的漂亮工作.假设流形具有极大欧氏体积增长和有界曲率,Chen-Tang-Zhu[87](对复二维) 以及 Ni[381](对所有高维) 证明流形可以紧化为一个复簇.去年,Albert Chau 和谭联辉[80] 终于在假设极大欧氏体积增长和有界曲率的条件下,解决了这个猜测.他们用到了倪磊[381] 的一个重要引理.这个引理证明了我的一个猜想(见[549] 或[381] 的引言部分).(即是说,极大体积增长蕴含数量曲率在平均意义下二次衰减)

虽然我们在正曲率情形看到了很大的进展,可是当凯勒流形是单连通,完备和强负曲率时,只有很少的结果.有一个猜想说,它是 $\mathbf{C}^n$ 中的有界域.(有人告诉我 Kodaira 考虑过类似的问题,可是我没有找到相应的出处) 主要的问题就是构造有界全纯函数.

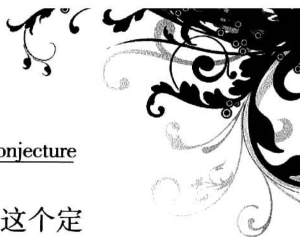
构造有界全纯函数的困难在于,Hörmander 的 L^2 方法的基本原理源于 Kodaria 消没定理.从加权 L^2 空间的函数过渡到有界函数是很困难的.另一方面,我证明了非负 Ricci 曲率完备流形上不存在非平凡有界全纯函数[545].

如果流形是一个紧致凯勒流形 M 的万有覆盖,其中 M 具有到一个亏格大于 1 的紧致黎曼面的非平凡拓扑映射,那么就可以用 Jost-Yau[256] 的证明方法来构造有界全纯函数.特别的,如果 M 有一个到亏格大于 1 的黎曼面的乘积的拓扑满射,并且具有非平凡拓扑度,那么万有覆盖应该很可能是一个有界域.

当然,这种构造是基于全纯函数是调和函数的事实.曲率的某些刚性迫使逆命题也成立.对函数来说,目标空间没有拓扑,刚性一般不成立.有界调和函数不能通过解 Dirichlet 问题来构造,除非加上更多的边界数据.比如边界有好的 CR 结构.事实上,对 $\mathbf{C}^n$ 中的奇数维子流形,如果每点的切空间都有极大复线性子空间,那么 Harvey-Lawson[222, 223] 证明了它们是复子流形的边界.不幸的是,一个具有有界负曲率的完备单连通流形的边界不是光滑的.定义奇异边界上的 CR 结构是很有意义的.我们提一下 Kuranish [280, 281, 282] 关于抽象 CR 结构嵌入的著名工作.

一类非常重要的丛是旋量(spinors) 丛.作用在旋量上的 Dirac 算子是一个主要的几何算子.Atiyah-Singer 是最早研究其几何性质的数学家,通过细致研究 Dirac 算子,他们证明了著名的指标定理[15].关于正数量曲率自旋(spin) 流形上调和旋量的 Lichnerowic [318] 消没定理给出了相当强的信息. 通过 Atiyah-Singer 指标定理,它给出了 $\hat{A}$ 亏格和正数量曲率自旋流形的 α 不变量的消没定理. 这一方法后来被 Hitchin [229] (对 K3 曲面) 以及





Gromov-Lawson[204,205]加以改进.对四维流形的spinc结构有效地应用这个定理在Seiberg-Witten理论中非常重要.Lawson-Yau [285]结合应用Lichnerowicz和Hitchin的工作证明了一大类光滑流形不存在非阿贝尔的光滑群作用.利用模形式,刘克峰[331]证明了与Lawson-Yau定理类似的环路空间上Witten亏格的消没定理.

在Schoen-Yau [425, 428]以及Gromov-Lawson [204, 205]的拓扑手术的结果基础上,我们相信Lichnerowicz定理的一个合理的逆成立.主要的结论是球面上余维数大于等于3的拓扑手术保持正数量曲率度量.一旦几何手术可行,配边理论中的标准结果使得我们可以对正数量曲率单连通流形推导一些存在性的结果.最好的工作属于Stolz [466],他对维数大于4的单连通流形给出了完全的解答.我也建议应该在具有正数量曲率共形边界的渐近双曲流形上施行手术.这个建议与最近Witten-Yau[534]关于共形边界连通性的工作有关.

研究正数量曲率度量是了解广义相对论中正质量猜想的第一个重要步骤.Schoen-Yau [426,430]用极小曲面的想法给出了第一个证明.三年以后,Witten [527]用调和旋量给出了另一个证明.Schoen-Yau接着把他们的证明推广到正数量曲率高维流形[428].后来Gromov-Lawson [204, 205]注意到Lichnerowicz定理可以与基本群结合,给出了正数量曲率度量的拓扑阻碍.这些工作与Novikov猜测有关.Lusztig[342], Rosenberg [404], Weinberger [524]与余国良[557]对这个猜测作出了贡献.

除了在论证Minkowski时空稳定性中的作用,正质量猜想被Schoen [419]用来证明Yamabe问题,其中Trudinger [507]和Aubin [16]也作出了重要的贡献.

Korevaar-Mazzeo-Pacard-Schoen [276]发展了有用的方法来理解Nirenberg的关于指定曲率的问题.接着是Chen-Lin [93], Chang-Gursky-Yang [79].可是,Kazdan和Warner的可积性条件还未充分理解.

流形的谱及其整体拓扑的关系是非常深刻的.Atiyah-Bott-Hirzebruch-Singer是最早研究它们关系的学者.Atiyah-Bott-Patodi[9]把热核展开与局部指标定理的证明联系起来.Atiyah-Patodi-Singer [12, 13, 14]首创了谱流的研究,给出了奇数维流形上的重要的全局谱不变量.这些全局不变量成为了Atiyah-Donnelly-Singer[10]和Mark Stern [465]所发展起来的L^2指标定理的边界项.Witten[528, 529]将超对称与de Rham复形的解析形变引入Morse理论,从而揭示了整体几何与理论物理的崭新联系.Witten的工作由Demailly [139], Bismut-Zhang [41,42]作了推广,用来研究全纯Morse不等式和解析扭矩.Novikov [383]研究了1形式的Morse理论.Witten有关Morse理论的工作引发了Floer(见[164, 165,166])的工作,他利用Floer上同调的思想来证明流形在高阶

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

同伦群消没时的 Arnold 猜想. Floer 的理论与三维流形纽结理论有关(通过 Chern-Simons 理论[109]). Atiyah, Donaldson, Taubes, Dan Freed, P. Braam 和其他许多学者(见[6, 479, 53, 167])都对这个课题作出了贡献. Pukaya-Ono [177], Oh [384], Kontsevich [272], Hofer-Wysocki-Zehnder[237], G. Liu-Tian [330],都研究了 Floer 理论与辛流形的关系. Arnold 猜想得到了部分证明. 可是完全满意的证明还未见到.

我们也要提一下 Cheeger [82] 和 Müller [375] 的重要工作,他们验证了关于解析扭矩与流形组合结构之间关系的 Ray-Singer 猜想. Ray-Singer [406] 关于全纯扭矩的基本想法正在 Vafa 等人[37] 的漂亮工作中得到蓬勃的发展. 解析扭矩的更多工作由 Quillen, Todorov, Kontsevich, Borcherds, Bismut, Lott,张伟平和陆志勤等作出(见[40]及其参考文献,[253], [47, 48]). Atiyah-Bott-Patodi [9] 的指标定理的局部情形由 Bismut [39] 的复杂工作推广到一族椭圆算子的情形. (局部指标定理的证明可以追溯到 Mckean-Singer [352] 的基本工作)Quillen 的工作在这个方向上是非常重要的(见[402, 403]). Witten [530], Bott-Taubes [50], Taubes[478],刘克峰[332] 以及 M. Hopkins[240] 在椭圆亏格的工作建立了拓扑与模形式之间的桥梁.

我们不应忘记,在 20 世纪 70 年代的时候,Bot[49] 和 Atiyah-Bott[8] 发展了等变上同调的局部化公式. 这已经发展成为了一个重要的工具. 例如,它被刘克峰及其合作者以最强有力的形式用于多个领域的研究:镜像原理(Lian-Liu-Yau [314, 315, 316, 317]),拓扑顶点(Li-Liu-Liu-Zhou [295]) 等. 最早将局部化应用于计数几何的是 Kontsevich [271],以及后来的 Givental [186] 和 Lian-Liu-Yau [314] 的独立工作. 这些工作解决了 Candelas 等人[70] 基于镜像对称的组合等式的猜想,这是共形场论为数学中古典问题提供灵感的范例.

全纯线丛的全纯截面在代数几何中始终是重要的研究对象. Riemann-Roch 公式与消没定理结合给出了线丛截面存在性的非常有用的结果. Kodaira 嵌入定理[269],即每个 Hodge 流形是射影的,开创了凯勒流形的全纯嵌入理论. 例如,Hirzebruch-Kodaira [228] 证明了每个与射影空间微分同胚的奇数维凯勒流形与射影空间双全纯同构.

给定线丛的全纯截面的一个规范正交基,我们就可以把它嵌入到射影空间中. 诱导度量是与线丛相配的 Bergman 度量. 注意 Bergman 度量的原始定义用到了典则线丛与 L^2 全纯截面.

在理解流形稳定性与 Kähler-Einstein 度量存在性的关系过程中,我[548] 建议,如果我们允许线丛的幂可以任意大,那么每个具有整凯勒类的凯勒度量可以用伯格曼度量来逼近. 遵循着 Siu-Yau[457] 文章中的想法,田刚[497] 在我的指导下在博士论文中证明了 $\mathbf{C}^2$ 收敛性. 我的另一个学生阮卫东[405] 在博士论





文中证明了 C^{∞} 收敛性.接着是 Lu[338], Zelditch[558] 和 Catlin[77] 的工作,他们注意到核函数的渐近展开可以从 Szegö 核的标准展开得到,这可以追溯到 Fefferman [160] 和 Boutet de Monvel-Sjöstrand[51] 关于凯勒流形上圆周丛的工作.

Kodaira 利用正线丛的充分高次幂嵌入 Hodge 流形的证明从幂次来说并不是有效的. Matsusaka[348] 以及后来的 Kollár [270], Siu [454] 的工作使得幂次成为有效的. Demailly [140, 141] 和 Siu [454, 456] 给出的重要定理几乎解决了著名的 Fujita 猜想[174](也见 Ein 和 Lazarsfeld [157]).萧荫堂的有效的方法也导致了代数流形多重亏格形变不变性的证明[455].

2 流形间的映射与几何结构的刚性

另一个推广函数概念的方向是研究流形间的映射.我们可以定义流形间映射的能量,其临界点称为调和映照. Eells-Sampson [156] 最早在像流形具有非正曲率的条件下证明了调和映照在同伦类中的存在性.

当我做博士论文的关于流形非正曲率与基本群的关系时,我意识到有可能用调和映照来改进我的论文中的某些定理.我也同时注意到调和映照在几何刚性问题研究中的重要性.当时,最著名的定理是 Mostow 刚性定理[376]. 1976 年,我证明了卡拉比猜想,并且应用 Kähler-Einstein 度量和 Mostow 刚性定理证明了球的商空间上复结构的唯一性[542].因此我也建议用调和映照来证明强负曲率凯勒流形上复结构的刚性.我向萧荫堂提出了我的建议,他在像流形满足更强的曲率条件下证明了我的想法[451].

Jost-Yau [257], Mok-Siu-Yeung [370] 还得到了关于 Margulis 超刚性定理 [346] 证明的更进一步的结果,改进了 Corlette [133] 的一个早期的结果,他证明了某些秩 —— 局部对称空间的超刚性.对复球商空间的超刚性的完全了解还未及理清.我们需要找到反映流形的基本结构的更多调和映照结构.此想法基于 Matsushima [350] 的证明方法,是 Calabi 所建议的.(这是 Calabi 在普林斯顿高等研究所几何特别年上讨论的专题)

我们也考虑离散化的调和映照. 当像流形是度量空间时, 有 Gromov-Schoen[206], Korevaar-Schoen [277] 和 Jost [255] 的工作.当始流形是单纯复形,则有 Ballmann-$\bar{S}$wwiatkowski [23] 和王慕道[517, 518] 的文章,他们引入了建物(building)概念的推广.他们还推广了 H. Garland [179] 关于 P 进制建物上同调群的消没定理.

应用重心的概念, Besson-Courtois-Gallot [38] 给出了秩一局部对称空间的度量刚性定理.还有 Gerver-Nadirashvili [184], Pestov-Uhlmann [392] 漂亮的度

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

量几何刚性反问题，即当我们知道边界上任意两点间的距离函数，我们可以求出黎曼度量.

黎曼面之间的调和映照有很好的性质.Morrey 最早解决了到任意黎曼流形的能量极小调和映照的狄利克雷问题.

另一个重大的突破是 Sacks-Uhlenbeck [413] 在 1978 年作出的，他们构造了黎曼流形中表示第二同伦群的极小球面，用到了调和映照在孤立点处的一个漂亮的延拓定理.通过细致改进他们的方法，Siu-Yau[458] 研究了调和映照的冒泡(bubbling) 过程，并用来证明了一定曲率条件下全纯球面的存在性.作为一个应用，他们证明了 Frankel 的著名猜想，即正双截曲率凯勒流形是 $\mathbf{C}P^n$.(这也被 Mori[374] 独立证明)

Gromov[202] 于是注意到近复结构的伪全纯曲线可以类似地用来证明 $\mathbf{C}P^n$ 上辛结构的刚性.上面提到的冒泡过程被 Ye[555] 和 Parker-Wolfson [388] 进一步改进来给出伪全纯映照模空间的紧化.基于这些思想，Kontsevich [271] 引入了稳定映射及其紧化的概念.

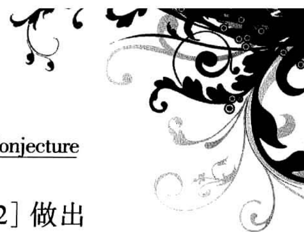
Gromov-Witten 不变量以及量子上同调的形式化定义是建立在这些发展和物理学家思想的基础上的.例如，量子上同调是由 Vafa(见[513]) 与他的合作者倡导研究的(这个名字是由 Greene 和我建议的).量子上同调的结合律是由四位物理学家 W-D-V-V[532,142] 提出的.数学上的验证(由 Ruan [411] 与 Ruan-Tian [412] 作出) 沿用了物理学家的黏合的思想.Ruan-Tian 用到了 Taubes[477] 的想法.可惜重要之处被忽视了.A. Zinger [563, 564] 最近完成了这些论证.

与 Donaldson 理论非常类似，我们需要在稳定映射模空间中引入虚拟闭链的思想.这些概念的代数理论比辛几何的情形要更加深刻，与代数几何发展也更密切相关.主要的想法是李骏贡献的，他也对 Donaldson 理论的代数几何情形作出了重要贡献(见[292, 296]).(他还在相对 Gromov-Witten 不变量理论的代数情形作出了关键的贡献[293, 294])Li-Tian [297] 的辛几何情形遗漏了许多难点，最近大多被 A. Zinger [563, 564] 补上了.

Schoen-Yau[429] 在 1976 年研究了 L_1^2 映射对流形基本群作用的意义.这被用来证明在基本群上具有给定作用的调和映照的存在性.最近，林芳华进一步发展了这个想法[325]. 他详细研究了映射空间中的几何测度论(见[322, 324]).第二同伦群上的作用理解起来要困难得多.我想要存在一个非平凡作用在第二同伦群上的调和映照，只需这样的连续映射存在.

从 S^2 到另一个流形的映照的能量泛函并不能立即与 Morse 理论建立联系.可是 Micallef-Moore [362] 应用 Sacks-Uhlenbeck 的摄动方法已经足够给出正迷向曲率流形的一些结构定理.





高维调和映照正则性理论的主要工作是 Schoen-Uhlenbeck[421,422] 做出的. Leon Simon [446] 对调和映照与极小子簇的奇点理论作出了重大的贡献.此外还有林芳华[324] 的工作.

在 20 世纪 80 年代早期,谷超豪[208] 开创了调和映照的一个新的分支.他最早研究始流形是 Minkowski 空间的调和映射.这个新分支现在称为波映照理论.这门学科由 Christodoulou, Klainermann, Tao, Tataru 和 M.Struwe 做了详细研究(见[118, 265, 474, 475, 438]).

Uhlenbeck [511] 应用可积系统的技巧构造从 S^2 到 $U(n)$ 的调和映照. Bryant [57] 和 Hitchin [232] 应用旋量理论和谱曲线也能构造一些调和映照.受这些工作启发,Burstall, Ferus, Pedit 和 Pinkall [62] 构造了从环面到任意紧对称空间的调和映照.在一系列的文章中,Terng 与 Uhlenbeck [490, 491] 用环路群因子分解,解决了逆散射问题,并且对孤立子方程构造了 Backlund变换,包括从 $\mathbf{R}^{1,1}$ 到 Hermitian 对称空间的 Schrödinger 映照.

3 由变分原理定义的子流形

给定黎曼面 Σ 上的一个共形结构,从 Σ 到一个固定黎曼流形的调和映照可以在给定的同伦类中使得能量极小化.可是,它不一定是共形的,也不一定是极小曲面.为了得到一个极小曲面,我们也需要变动 Σ 上的共形结构.既然曲面上共形结构的空间不是紧致的,我们需要保证极小总是可以取到.

如果映射 f 诱导流形基本群的单射,Schoen-Yau 通过应用 Linda Keen 的一个定理证明了能量在模空间上是逆紧的.(他们发展了映射的拓扑理论) 从而他们证明了不可压缩极小曲面的存在性[425].作为这个论证的一个副产品,我们有可能找到 Teichmüller 空间的一个好的穷竭函数. Michael Wolf[535] 用调和映照给出了 Teichmüller 空间的一个紧化,并且证明它等价于 Thurston 紧化. S.Wolpert 详细研究了 Weil-Petersson 度量的行为(见 Wolpert 的综述文章[536]). Royden [410] 证明 Teichmüller 度量就是 Kobayashi 度量. C.McMullen [377] 引入了模空间上新的凯勒度量,可以用来证明模空间在 Gromov 意义下是双曲的. Teichmüller 空间中各种内蕴度量的比较是一个重要的问题[548].最近被 Liu-Sun-Yau [333, 334] 所解决.事实上 Liu-Sun-Yau 引入了具有有界负曲率和几何的新度量,发现了模空间对数余切丛的稳定性. Schoen-Yau[424] 还证明了次数为 1 的调和映照是一一的,如果像曲面具有非正曲率.这种类型的结果只对二维曲面成立.

欧氏空间中极小曲面理论有一段悠久而丰富的历史.最近 Meeks, Osserman, Lawson, Gulliver, White, Hildebrandt, Nadirashvili 和其他学者(见

Colding和Minicozzi的综述文章[131]中的参考文献)关于欧氏空间嵌入极小曲面有很多漂亮的工作.他们很接近于给出完备嵌入极小曲面的分类,对有界域中完备极小曲面有了更好的了解.例如,Meeks-Rosenberg [353] 证明平面和螺旋面是 $\mathbf{R}^3$ 中仅有的逆紧嵌入单连通极小曲面.

在1978年开始的一系列文章中,Meeks-Yau [355, 356, 357, 358] 解决了一个重要猜想,即如果边界曲线是凸边界的子集,那么Douglas关于Plateau问题的解是嵌入的.(注意Osserman[386]早就解决了Douglas解中分支点的非存在性这个古老的问题.而Gulliver [213] 证明了false分支点的存在性)我们利用极小曲面的面积极小化的事实证明这些曲面在群作用下是等变的.等变嵌入曲面在流形的有限群作用中扮演了重要角色.结合Thurston的一个定理,我们证明了对作用在球面上的循环群的Smith猜想[554]:不动点集不是一条扭结曲线.

Meeks-Yau在固定曲面亏格的条件下应用Douglas-Morrey对于Plateau问题的解. 如果亏格可以任意大,则极小化面积是个困难的问题. 这已经被Hardt-Simon[221]解决.在另一个方向,Almgren-Simon[3]成功地极小化了欧氏空间中的嵌入圆盘.Meeks-Simons-Yau [354] 应用这一技巧证明了包围伪球的嵌入极小曲面的存在性.这个重要的定理可以用来证明不可约三维流形的万有覆盖也是不可约的.他们还给出了高亏格嵌入极小曲面存在的条件.拓扑学家Freedman-Hass-Scott [169] 继续了这方面的工作.Pitts [395] 用多形体的极小-极大方法证明了嵌入极小曲面的存在性.Simon-Smith在未发表的文章中证明了具有任意度量的三维球面上嵌入极小-极大球面的存在性.Jost[254]用改进的方法找到了四个极小-极大球面.Pitts-Rubinstein(见[396])继续研究了这种极小-极大曲面.三维流形拓扑学家用Meeks-Yau的思想处理组合型的极小曲面,并给出了在三维流形拓扑中的应用.

最近Colding和Minicozzi [126, 127, 128, 129] 在有关极小曲面迭片结构和面积无限大的极小曲面估计方面,取得了显著的成果,很可能会在三维流形拓扑中找到重要应用.事实上,他们[130]对Hamilton的Ricci流的消没时间估计作出了贡献.

高维极小簇在几何中非常重要.有Federer-Fleming [159], Almgren [2] 和Allard[1]等人的工作.Bernstein猜想(即极小图是线性的)的证明对极小簇理论发展起了很大的推动作用.Bombieri, De Giorgi和Giusti [46] 发现了Bernstein问题的著名反例.这激起了许多人对于面积极小化锥(因为极小图必是面积极小化的)的兴趣.Schoen-Simon-Yau [420] 找到了一个证明低维Bernstein问题的不同途径.这篇有关稳定极小超曲面的文章极大推动了稳定超曲面的曲率估计.还有L. Simon, Caffarelli和Hardt[64]利用形变稳定极小锥构造极小超曲面的工作.最近,N. Wick-ramasekera [525, 526] 在稳定极小(分支)超曲面上作了





一些深刻的工作,推广了 Schoen-Simon-Yau 的结论.

高余维数的最好正则性结果是 Almgren 得到的,他证明了对面积极小簇,奇异集的余维数至少为 2.这样的结果怎样应用到几何上还不清楚.也许最重要的问题是 Hodge 猜想,即射影流形中一个(p,p)型上同调类的倍数是否可以用一个代数闭链来表示.Lawson 作了一个尝试,应用了 J.Simons 的阻碍性定理 —— 复射影空间中的稳定子簇是代数的[283].怎样利用(p, p)型的假设是一个困难的问题.一般来说,代数闭链不是有效的.这也让解析方法产生了很大的困难.King [263] 和 Shiffman[443] 关于复流(current) 工作也许会有所关联.

也许我们应该把 Hodge 猜想加以推广,包含进一般的(p,q)类,因为很可能 $\bigoplus_{i=-k}^{k} H^{p-i,p+i}$ 中的每个闭链都有理同调于极小子簇的和,使得这个簇的几乎每个点处的法空间都有一个 $P-k$ 维的复空间.

重要的一点也许是要求度量是典则的,如凯勒 - 爱因斯坦度量.

一个对偶的问题是如何用极小 Lagrangian 闭链表示同调类. 当流形是 Calabi-Yau 时,这些是特殊 Lagrangian 闭链.既然它们对偶于全纯闭链,那么也应该有一个 Hodge 猜测的类比.例如,当 $\dim_{\mathbf{C}} M = n$ 是奇数时,$\bigoplus_{i+j=n} H^{i,j}$ 中的每个元素如果它与凯勒类的卡(cup) 积为零, 则此元素应该可以用特殊 Lagrangian 闭链的有理数倍来表示.

一个相关的问题是,如果复向量丛的陈类是(p,p)型的,这个向量丛在加上另一个全纯向量丛以后是否容许一个复结构?如果上面 Hodge 猜测的推广成立,那么向量丛也应该有一个类似的推广.

由于 Taubes 的关于伪全纯曲线[486] 和反自对偶丛[476,477] 的存在性工作,这些问题在四维辛流形理论中取得了很大的成功.在凯勒曲面上,反自对偶联络是全纯向量丛的 Hermitian 联络.特别地,Taubes 给出了构造凯勒曲面上全纯向量丛的方法.

另一类重要的极小簇是 Calabi-Yau 流形中的特殊 Lagangian 闭链.这些闭链是 Harvey-Lawson [224] 在发展欧氏空间中的定标(calibrated) 几何时引入的.主要的工作是 Schoen-Wolfson [423],李莹英[288] 和 Butscher [63] 做出的.我们认为 Lagrangian 闭链与全纯丛对偶,而特殊 Lagrangian 闭链与 Hermitian-Yang-Mills 联络互为镜像.所以根据 Donaldson-Uhlenbeck-Yau 定理,这就与稳定性有关. Joyce 和 Thomas 讨论了 Lagrangian 闭链稳定性的概念. 在镜对称的基础上, Thomas-Yau [492] 建议用平均曲率流的方法从稳定的 Lagrangian 闭链出发去证特殊 Lagrangian 闭链的存在性.特殊 Lagrangian 闭链的存在性与平均曲率流的关系是非常有趣的问题. 参看王慕道[519, 520], Smoczyk [461] 和 Smoczyk-Wang[462]. 镜对称几何被 Strominger-Yau-Zaslow [468] 用特殊 Lagrangian 环面的纤维化来解释镜对称几何.还存在具有其他特殊和乐群的流

形,相应的有类似的定标子流形.Conan Leung 对这种流形及其镜像的研究作出了贡献(见[290, 291]).

除了极小曲面,重要的曲面还包括常平均曲率曲面与平均曲率范数达到极小值点的曲面.知道这种曲面在非负数量曲率三维流形中的存在性是重要的问题.

极小球面的存在性与黑洞的存在性有关.最有效的方法是 Schoen-Yau [432] 发展的,而这要用他们[427] 证明的 Jang 方程存在性定理.发展证明稳定极小球面的存在性定理的方法是很有意义的.Hawking 质量的极值与平均曲率 L^2 范数的极小值点有关.其存在性与性质都还没有充分了解.

对常平均曲率曲面,我们有稳定性的概念.(固定它所包围的体积,那么面积的第二变分是非负的) 利用我与李伟光在特征值方面的工作,Christodoulou 和我[119] 证明了这些曲面的 Hawking 质量是正的.这个事实被 Huisken 和我[248]用来证明具有正质量的三维渐近平坦流形的"常平均曲率球面叶状结构"的存在性与唯一性.(我们从 1986 年开始这项研究.Ye 研究了我们的工作,在各种不同条件下证明了类似叶状结构的存在性,见[556]) 十年以后,田刚和庆杰[401] 基本上用我们的方法重新证明了我们的定理.

这种叶状结构被 Huisken 和我[248] 用来给出无穷远处的一个典则坐标系.这就定义了重心的概念,其中可以找到广义相对论的重要性质.最值得注意的是全线性动量等于全质量乘上质心的速度.我们希望找到广义相对论中张量沿着这些典则曲面的好的渐近性质.我们也想在这个重心概念的基础上找到角动量的一个好的定义,使得全质量可以控制角动量的平方范数.一些初步的进展已经由我的学生张晓[561] 作出,可是他的定义没有好的物理背景,目前还没有找到比较好的应用.

在 Huisken-Yau [248] 的文章中,叶状结构的唯一性被改进为包围三维流形上一个紧集的闭超曲面的唯一性.这个紧集是用[248] 中超曲面的平均曲率定义的.作为一个很次要的注记,Huisken-Yau 提出这个唯一性应该也与平均曲率的选择无关.虽然这个改进对于单个的超曲面本身有点意思,但是对于如上广义相对论中的应用毫无关系.特别的,它与质心的定义无关.这个 Huisken-Yau 的建议被 Qing-Tian [401] 证明是正确的,他们用的基本上是与 Huisken-Yau 同样的论证,以及一个简单的积分公式.

用满足各种性质(如常高斯曲率) 的曲面的叶状结构来研究广义相对论中的三维流形的想法也由 R. Bartnik [30] 作了发展.他的关于拟球面叶状结构的想法为一大类正数量曲率度量给出了很好的参数化.

这些想法中有一些被 Shi-Tam [442] 用来研究与包围三维正数量曲率流形的球面相关的量.这些量后来被意识到就是 Brown-York [56] 的拟局部质量.同





时,刘秋菊和我[328, 329] 给出了广义相对论中一般时空的拟局部质量的新定义,其中用到了 Shi-Tam 的一些想法.我和王慕道的进一步的工作表明了理解双曲空间形式中曲面的重要性.

我对拟局部质量的兴趣可以追溯到我与 Schoen [432] 关于质量浓缩导致黑洞存在的文章.我们很希望找到一个拟局部质量,包含了物质的效应与重力的非线性效应.我们有可能证明当这一质量大于面积平方根的一个常数倍,则黑洞形成.这还未得到证明.

Huisken-Yau 的证明方法依赖于平均曲率.还有另一种称为反平均曲率的流,是 Geroch [183] 为了理解联系质量与黑洞面积的 Penrose 猜想而提出的.当数量曲率非负时,这些证明的步骤由 Huisken-Ilmane [245] 给出.后来 H. Bray [55] 给出了一个不同的证明.多形体(varifold) 的平均曲率流是 Brakke [54] 首先研究的.许多学者研究了水平集方法:S. Osher, L. Evans, Giga 等(见[385, 158, 96]).Huisken[243, 244] 首先在曲面为凸时做了重要的工作.他最近与 Sinestrati [246,247] 关于平均凸曲面的工作非常出色,给出了平均曲率流奇点结构的非常不错的理解.

自然的高阶椭圆问题很难处理.仿射极小曲面和 Willmore 曲面就是这样的例子.L. Simon [445] 对 Willmore 曲面的正则性作出了重要的贡献.

大多数椭圆方程有一个抛物的版本.可是要在流形上找到一个能应用极值原理的合适的量不是一件容易的事情.除了 Li-Yau [309] 的工作,我希望 Cheng-Yau 关于极大片和仿射球面的梯度估计有一个抛物的版本.另一方面,凯勒度量的 Calabi 流(见 Chang [78] 的综述文章)还未见到太多进展.这些是高阶的问题,极大值原理并不是那么有效.Chrúsciel [120] 对黎曼面情形作出了重要的贡献.

Einstein 方程的动力学性质是一个非常困难的课题.许多学者考虑了 Cauchy 问题:Lichnerowicz, Choquet-Bruhat, York, Moncrief, Friedrich, Christodoulou, Klainerman, H. Lindblad, Dafermos(见[319], [115], [114], [172],[116],[264],[326],[136]).可是整体的行为还远未理清.主要的未解决问题是陈述和证明 Penrose 关于宇宙监督(cosmic censorship) 的基本问题.我向 Klainerman 建议考虑 Einstein 方程组的小初始值解.Minskowski 时空稳定性的处理被 Christodoulou-Klainerman [117] 在较小下降条件下完成.最近 Lindblad 和 Rodnianski [326] 给出了一个更简单的证明.几年前,N. Zipser 在哈佛博士论文中把 Maxwell 方程加到重力上,仍然证明了 Minkowski 时空的稳定性. Schwarzschild 或 Kerr 解的稳定性还远未了解清楚.Finster-Kamran-Smoller-Yau [162] 研究了这种背景下 Dirac 粒子的衰减性质.这项工作非常深刻,表明了这些经典时空的稳定性.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

关于常平均曲率类空超曲面有大量文献.它们定义的叶状结构可以给出时空上有趣的动力学性质.这些曲面即使对于 $\mathbf{R}^{3,1}$ 也是很有趣的.Treiberges [506] 对它作了详细研究.Li, Choi-Treiberg [111] 和温有恒[515] 注意到这种曲面的高斯映照给出了到圆盘的调和映照的很好例子.最近 Fisher 和 Moncrief 应用它们研究了 2 + 1 维 Einstein 方程的演化.

4 从与流形上几何结构的构造

在给定流形上构造几何结构是一个基本的问题.一般而言,保持这个几何结构不变的拓扑等价群应该是一个特殊群.除了辛结构,这些群总是有限维的.当几何结构在等价意义下是唯一的,那么它可以用来给出拓扑结构的关键信息.

对特殊几何结构的研究可以追溯到 Sophus Lie, Klein 和 Cartan.在大多数情况下,我们希望平移向量来定义和乐群的概念.当和乐群不是紧的,有射影平坦结构,仿射平坦结构和共形平坦结构这些例子.要决定哪些拓扑流形容许这些结构不是一件容易的事.既然结构是平坦的,就有唯一连续延拓的性质,所以我们可以构造一个发展映射,从流形的一个合理覆盖分别到实射影空间,仿射空间和球面.这个映射给出了流形基本群到实射影群,特殊线性群和 Möbius 群的表示.这个和乐表示给出了几何结构的大量信息.但是这个映射一般不是单的.如果它是单射,那么流形可以看做区域在对应李群的一个离散子群作用下的商空间.在这种情况下,可以更好地从偏微分方程角度看待流形,离散子群也起了重要的作用.比如,如果一个射影平坦流形可以射影嵌入为一个有界域,Cheng-Yau [104] 从实的 Monge-Ampère 方程构造了一个典则度量,推广了 Hilbert 度量.当流形是二维的,有王长平[516] 和 Loftin [335] 的关于如何把这个度量与共形结构,以及典则丛三次幂的一个全纯截面结合起来的工作.这是一个漂亮的理论,与第一节中提到的双曲仿射球面有关.

关于黎曼面上射影结构的几何分解和模空间,有 Choi, Goldman(见 Choi 和 Goldman [113] 的文章中的参考文献) 以及 Hitchin [233] 等人的基本工作.把这些推广到三维或四维流形也许会很有意义.

决定哪些流形容许仿射结构是一个困难的问题.一个公开的问题是这种空间的欧拉数是否为零.Sullivan[470] 在这个问题上已经作出了重要的贡献.Goldman [187] 也找到了三维流形上用基本群表示的拓扑限制.困难来自于不存在有用的与仿射结构相容的度量.这促使 Cheng-Yau [105] 定义了仿射凯勒度量的概念.

当郑绍远和我考虑仿射凯勒度量概念的时候,我们认为它是凯勒度量的自





然类比.可是,紧致非奇异的例子并不多见.Strominger-Yau-Zaslow [468] 建议通过用特殊 Lagrangian 环面构造 Calabi-Yau 流形的商空间的办法来构造镜流形.在大凯勒类的极限,Hitchin[234] 指出商空间容许一个与仿射凯勒结构相容的自然仿射结构.可是一般来说,这种结构存在奇点.一个深刻的问题是理解奇异性的类型以及如何从这种奇异结构构造 Calabi-Yau 流形.Loftin-Yau-Zaslow [336] 研究了"Y"奇点的结构.我们希望可以找到紧致流形上沿给定的余二维分层子流形为奇点的仿射结构.

构造共形平坦流形也是一个有趣的问题.与射影平坦或仿射平坦流形类似,共形平坦结构的存在性有一个来自于 Pontrjagin 类的曲率表示的拓扑障碍.更深刻的问题与对基本群和发展映射的理解有关.当结构容许一个具有正数量曲率的共形度量.Schoen-Yau[433] 证明了非常有趣的定理,即发展映射是单的.所以这种流形必是 S^n 中一个区域在 Möbius 变换群的离散子群作用下的商空间.分类这种流形是很有意思的问题.在这点上,Schoen [419] 所解决的 Yamabe 问题提供了常数量曲率的共形度量.我们希望能够用这种度量来控制共形结构.然而,这种度量不是唯一的.于是对常数量曲率共形结构模空间的更深入理解变得尤为重要.

很奇怪的是,丛理论在黎曼几何中得到了广泛的使用,可是却没有用在这些几何中.我们可以构造实射影空间丛、仿射丛或球面丛,方法是通过把坐标图映射到对应的模型空间(也许维数不同),然后把目标模型空间黏合起来得到自然的丛.我们也可以研究相应的 Chern-Simons 形式[109].

几何学的一个重要目标,就是构造一个给定拓扑下的典则度量.除了二维的单值化定理,这方面仅有的出色工作是 Thurston 的几何化猜想(见[495]).

Thurston 应用黎曼面的想法,Haken [214] 关于三维流形的工作,Mostow 刚性[376] 以及李群的离散群论,建立了他的纲领.他假设三维流形中不可压缩曲面存在.当 Hamilton [215] 在 Ricci 流形取得了初步的成果,我立刻建议他(大约 1981 年)用他的 Ricci 流来分解流形.他把 Li-Yau [309] 的理论推广到 Ricci 流 [216, 217],以及他在 1996 年[218] 关于流形分解的开创性工作是这个纲领的里程碑.Perelman最近的想法[389, 390] 正是建立在这两项工作的基础上,并且更加的深入.详细的讨论在过去两年中由朱熹平、曹怀东、Hamilton 和 Huisken 作出.我们有理由相信这将最终导致 Thurston 纲领的证明.Hamilton 和 Perelman 的理论应该看做过去三十年中几何分析的最大成就.这段时期发展起来的几何分析思想几乎都用到了.

拓扑学家用双曲度量给出三维流形的不变量.Thurston [494] 注意到双曲度量的体积是一个重要的拓扑不变量.模整定义的 Chern-Simon [109] 不变量可以看做这种流形的相位.这些不变量后来出现在 Wiitten 关于 2 + 1 维的重力理

论中[530]，S. Gukov [212] 把他们与纽结理论的基本问题联系起来.

Thurston，Hamilton，Perelman 等人的主要成就是创造了三维流形上的典则结构.目前对四维流形我们甚至还不能猜测它的典则结构.虽然 Donaldson 不变量和 Seiberg-Witten 不变量已经取得了重大的进展.Taubes [480] 对一大类一般四维流形证明了自对偶度量存在性.可是它们的模空间和拓扑含义目前都还不明朗.在 Taubes 的构造之前,Donaldson-Friedman [151]，LeBrun [286] 用扭子(twistor) 理论和代数几何的方法在 $\mathbf{C}P^2$ 的连通和上构造了这种自对偶的结构.

Atiyah-Hitchin-Singer [11] 为自对偶联络奠定了代数与几何的基础，而 Uhlenbeck[509，510] 奠定了分析的基础.她建立了 Yang-Mills 联络的可去奇点定理和紧性定理.这些定理导致了 Taubes[477] 和 Donaldson [143] 的基本工作，而这些工作极大变革了四维流形拓扑学.

Donaldson 理论的出发点是四维流形的反自对偶联络.类似地,在凯勒流形的全纯向量丛上有 Hermitian Yang-Mills 联络的概念.Donaldson [144](凯勒曲面) 和 Uhlenbeck-Yau(一般凯勒流形)[512] 证明了 Hermitian Yang-Mills 联络的存在性等价于丛的稳定性. 这一定理后来被 C. Simpson [447] 利用 Hitchin [231] 的想法推广到具有 Higgs 场的丛.这在 Hodge 结构的变化理论中有深刻应用[448,449]. Wentworth 和 Daskalopoulos(见[138]) 研究了这一理论在曲线上向量丛模空间的情形.Li-Yau [298] 把 Hermitian Yang-Mills 联络的存在性推广到非凯勒流形上去. (Buchdahl [60] 后来对复曲面做了同样的工作)Li-Yau-Zheng [300] 用这个结果给出了 VII_0 类曲面的 Bogomolov 定理的完整证明.我希望 Donaldson-Uhlenbeck-Yau 理论在代数几何能找到更多的应用.

D.Gieseker [185] 对丛的模空间发展了几何不变量理论，引入了丛的 Gieseker稳定性.Conan Leung [289] 在我的指导下,在他的博士论文中引入了与 D. Gieseker稳定性相对应的分析概念(某一 PDE 的解).虽然这个分析概念的定义是自然的,仍有一个分析问题需待解决.(在丛的曲率是一致有界的假定下,梁证明了两个概念的等价性)

Bartolomeis-Tian[32] 尝试把 Yang-Mills 理论推广到辛流形,以及 Tian [500] 试图把它推广到具有特殊和乐群的流形.证明还不是完备的,有待完善.

对流形上的一个给定的自然结构,我们可以将方程线性化得到切丛上的一个自然联络.通常我们总是得到 Yang-Mills 联络,以及一些由原始结构的和乐群给出的额外结构.令人好奇的是,是否可以找到一个迭代的过程来找到有趣的度量.无论如何,我们可以在丛理论与度量理论之间找到类似的性质.丛的稳定性概念在全纯范畴中得到了比较好的理解.我相信对每个自然的几何结构,都应该有一个稳定性的概念.Donaldson [145] 通过推广 Atiyah-Bott [7] 关于黎曼面上丛的工作,用动量映射给出了稳定性的解释.对其他几何结构找到动量





映射是很有意义的.

最有趣的几何结构是凯勒结构.凯勒结构是由复结构与辛结构来定义的.复结构在二维时是非常刚性的,但在二维以上的流形上找复结构就会比较容易些.比如,Clemens [122] 和 Friedman [171] 提出了一种复手术的方法,即将余维大于一的具有负法丛的复子流形收缩成一点,再把收缩成的奇点光滑化,就可得到有趣的非凯勒复流形.(Lu-Tian [339] 就是按着这些想法构造例子的) 这些流形可以在连接 Calabi-Yau 流形方面起到重要作用,这是 M. Reid[407] 的猜测.为了进行更深刻的分析,给它们赋一个典则的度量很重要.(这种度量的一个候选就是 Fu-Yau [173] 研究的 Strominger 系统)

如同 Smith-Thomas-Yau [460] 所指出的那样,非凯勒流形的镜像可以是辛流形.他们通过倒转 Clemens-Friedman 的步骤,构造了一大类第一陈类为零的辛流形.李天军[311] 确定了第一陈类为零的四维辛流形的 Betti 数.在过去十年中,Gromov[202], Taubes [482, 483, 484,485], Donaldson [146, 147, 148] 和 Gompf[188] 等人对辛流形有了广泛的研究.这些工作是基于对伪全纯曲线和 Lefschetz 纤维化的理解.最基本的工作是围绕四维流形.主要的工具是 Seiberg-Witten[435, 436, 533] 理论和分析 Taubes 关于伪全纯曲线存在性及其计数的拓扑含义的工作是几何中最深刻的工作之一.在这项工作基础上,Taubes[482] 证明了一个古老的猜想,即标准 $\mathbf{C}P^2$ 上只有一个辛结构.我提出了如下问题:如果是同伦于 $\mathbf{C}P^2$ 的四维辛流形,是否辛同胚于标准 $\mathbf{C}P^2$?这问题到现在还没解决.(对应的复几何问题被我在[542] 中解决) 另一方面,在 Taubes[481] 工作基础上,李天军和刘艾克找到了容许正数量曲率四维流形的穿墙(wall crossing) 公式.接着刘艾克给出了这种流形的分类.(Stolz[446] 在 Schoen-Yau-Gromov-Lawson 的工作基础上得到的正数量曲率流形的手术在四维情形并不管用) 作为[312] 中一般穿墙公式的应用,李天军和刘艾克在[313] 中证明了在任意黎曼面的 S^2 丛上存在唯一的辛结构.他们用到了 D. McDuff [351] 的一个主要结果.

Strominger 系统[467] 解给出了超对称.第一陈类为零的辛流形上是否容许类似的结构是很有意思的问题.Fukaya 和 Oh [176] 对具有 Lagrangian 闭链的辛流形发展了一套理论.根据 Floer 理论,伪全纯圆盘是作为曲线运动的轨迹出现的.PaulSideal-Thomas[437] 和阮卫东讨论了与 Kontsevich 的同调镜猜测有关的 Fukaya 范畴.想知道的是 Fukaya 的理论能否对构造辛结构的典则度量有帮助.

Calabi-Yau 流形的构造是建立在容许凯勒结构的复结构存在性基础上的.一个基本的问题是一个复维数大于二的近复流形是否容许一个可积复结构.

一旦我们有了一个可积复结构,我们可以开始寻找具有特殊性质的 Hermitian 度量.如我们先前提到的那样,如果我们采用无扰的联络,那么从超

对称知道度量必须是凯勒度量,而流形必须是 Calabi-Yau 流形.如果不要求联络是无扰的,Strominger [467] 在不要求流形是情形下导出了一组反映超对称的方程.这是 Hermitian Yang-Mills 联络和 Hermitian 度量的耦合系统.二十年前由于我把自己限制在凯勒流形上,我没有导出这种系统.我的学生 Bartnik 与 Mckinnon [31] 合作得到了 Lorentzian 情形的耦合系统.他们找到了这种耦合系统的非奇异解.(Smooller-Wasserman-Yau-Mcleod [464] 和[463] 给出了数学上严格的证明)

李骏和我[299] 证明了 Strominger 系统在 Calabi-Yau 结构的一个邻域里可解.傅吉祥和我[173] 可以在不容许凯勒结构的复流形上对它求解.这类流形称之为平衡(balanced) 流形,由 M. Michelsohn[364] 首先研究的.这些流形可以用来解释弦论中流量(flux) 的问题(见[34,76]).

Calabi-Yau 结构在弦论中出现以前,就被我和其他学者用来解决代数几何中的重要问题.例如,Todorov [504]-Siu [453] 证明的 K3 曲面 Torelli 定理以及 Siu[452]-Todorov[504] 证明的 K3 曲面周期映照的满射性质,都是代数曲面的重要工作.Bogomolov[44]-Tian[496]-Todorov[505] 定理的证明也需要用这个度量.

在 1978 年国际数学家大会的报告中,我[544] 简述了非紧 Calabi-Yau 流形分类的提纲.一部分工作由 Tian-Yau [502, 503] 完成.在 1984 年期间,弦论学家迫切需要构造欧拉数等于 ± 6 的 Calabi-Yau 三重形(threefold).我[547] 构造了基本群为 $\mathbf{Z}_3$ 的这种流形.取两个三次曲面乘积中的(1,1) 次超曲面,其上有一个自然的 $\mathbf{Z}_3$ 作用,所得的商空间就是这种 Calabi-Yau 三重形.不久以后,我和田[501] 构造了更多的例子.

负数量曲率 Kähler-Einstein 度量的存在性是 Aubin [18] 和我[543] 独立证明的.我[542] 用它解决了代数几何中的经典问题,例如 $\mathbf{C}P^2$ 上复结构的唯一性[542],Miyaoka[365]-Yau[542] 陈数不等式,以及与 Shimura 簇双全纯同构的代数流形的刚性.一般情况下正数量曲率 Kähler-Einstein 度量的存在性还未被证明.可是,我对于 Calabi 猜测的证明已经提供了所有必要的估计,除了一些对于未知量的积分估计.这当然可以把它作为假定成立.我猜想这种积分估计与流形的稳定性有关.田[499] 称之为 K – 稳定性.Mabuchi 泛函[345] 把这个积分估计变得更加的内蕴,给出了这个问题自然的变分描述.萧荫堂指出田[498] 的关于二维曲面的工作还有漏洞.Nadal[378] 有关乘数层的工作给出了研究这个问题的有趣方法.

对正数量曲率 Kähler-Einstein 度量,有可能它们容许自同构的连续群.Mat-sushima [349] 最早注意到,这个群必是可约的.Futaki [178] 引入了一个有名的不变量,证明对这种流形它必定是零.在 20 世纪 80 年代我的讨论班上,我





建议 Futaki 的定理应该能够推广用来理解射影群在反典则高阶幂嵌入的流形上的作用,并且 Futaki 不变量应该与我[550] 的 Kähler-Einstein 流形与稳定性相联系的猜想相关.田刚曾问我当流形没有群作用是什么情况.我在班上向他解释,因为流形在射影空间里,它还是有射影群作用的影子,我们应该在射影群作用下把流形进行容许奇点的形变以得到更多信息.Futaki 不变量和流形稳定性的关系出现在最近 Donaldson [149, 150] 的工作中. 我们也提一下最近 Ross-Thomas [408, 409] 关于流形稳定性的有趣工作.还有许多用分析方法研究流形稳定性的文章.例如,Phong-Strum [393] 构造了某些退化 Monge-Ampére 方程的解,来得到凯勒位势空间中的测地线.他们[394] 也研究了 Kähler-Ricci 流的收敛性.

我提出的通过典则度量理解流形稳定性的建议被许多学者采纳.最好的结果是 Donaldson [149] 作出的.常数量曲率凯勒度量等价于第一陈形式是调和的.调和凯勒度量的唯一性定理是陈秀雄[95], Donaldson [149] 和 Mabuchi 分别在不同情形下给出的.(注意正数量曲率 Kähler-Einstein 度量的唯一性是 Bando-Mabuchi[26] 证明的) 一般情形的基于流形稳定性的调和凯勒度量的存在性还是未知的.在20世纪80年代中期我的讨论班上,我也讨论了调和凯勒流形与稳定性的关系:我的学生如田刚[497],罗华章[341] 和王晓伟[523] 的博士论文都与此有关.在他们之前,我的学生 Bando[25],曹怀东[71] 尝试用 Ricci 流研究这个问题.环簇上具有常数量曲率的调和凯勒度量是由 Donaldson [150] 首先研究的他建议通过解一个实 Monge-Ampére 方程来研究存在性问题.这一问题已由汪徐家与朱小华[522] 解决.在这之前,LeBrun 与他的合作者[262] 用扭子理论找到了对调和凯勒曲面的特殊构造.Bando 其实对具有调和第 i 陈形式的凯勒流形感兴趣.(对调和第 i 陈形式找到一个类似的稳定性是有趣的问题) 曹怀东在哥伦比亚大学讲授过凯勒度量的课程,当时张寿武正研究 Arakelov 几何.在他得知我的将流形稳定性与调和凯勒度量联系起来的猜测以后,他开始对将度量与流形的高度联系起来感兴趣.于是他与我的学生讨论流形稳定性与流形在复射影空间中嵌入所诱导的度量的关系[560].(后半部分与上面提到的 Bergman 度量有关) 他[560] 的关于 Chow 稳定性的工作对我的学生产生了很大的影响,包括田刚[32].张寿武文章的地位还未得到应有的重视.但他的工作对后来 Donaldson [149, 150] 的工作产生了重要的影响.

除了凯勒流形,还有具有特殊和乐群的流形.黎曼流形的和乐群是由 Berger [35] 分类的.最重要的是 $O(n)$, $U(n)$, $SU(n)$, G_2 和 $Spin(7)$.前面两个群分别对应于黎曼几何与凯勒几何.$SU(n)$ 对应于 Calabi-Yau 流形.G_2 流形是七维的,而 $Spin(7)$ 流形是八维的(假设它们是不可约的).后三类流形的 Ricci 曲率为零.值得一提的是,在我[543]1976 年证明 Calabi 猜测之前,还不知道任

何非平凡的紧致 Ricci 平坦流形.具有特殊和乐群的流形容许非平凡的平行旋量,它们对应于物理中的超对称.源自弦论中的想法对理解这些流形起了很大帮助.可是,在给定拓扑流形上如何构造这些结构,这个基本问题还未充分了解.在 G_2 和 $Spin(7)$ 情形,Bryant(见[58,59])首先作了研究.Joyce 构造了第一组紧致的例子[258, 259, 260].最近,Dai-Wang-Wei [137] 证明了具有平行旋量流形的稳定性.

Joyce 的漂亮的构造基于奇异摄动,类似于 Taubes [476] 关于反自对偶联络的构造.可是,这不是整体化的,无法给出 G_2 或 $Spin(7)$ 结构的好的参数化.这方面还有许多工作要做.Hitchin [235, 236] 关于三形式和四形式的工作也许在这些重要问题中有用.

我们也可以通过在具有特殊和乐群的流形上赋予一个奇异结构的办法得到新的几何结构. 比如, 如果我们要求一个度量锥容许 G_2, $Spin(7)$ 或 Calabi-Yau 结构,这个锥与单位球面的"交" 就是具有特殊结构的紧流形.锥上的度量给出了"交" 上有趣的 Einstein 度量.当锥是 Calabi-Yau 时,那么奇数维"交" 流形上的结构就称为 Sasakian 爱因斯坦度量.

在 Sasakian 爱因斯坦流形上存在一个自然的 Killing 场,称为 Reeb 向量场.如果它生成圆周作用,轨道空间就给出了正数量曲率的 Kähler-Einstein 流形.可是,Reeb 向量场不一定生成圆周作用, 由此,J. Sparks, Gauntlelt, Martelli, Waldram[180] 构造了许多有趣的非正则 Sasakian 爱因斯坦结构的显式例子.它们具有与共形场论有关的有趣的性质.对于拟正则的例子,有 Boyer, Galicki 和 Kállar[52] 的工作.这些过程给出了奇数维流形上的爱因斯坦度量的许多有趣例子.

一个固定拓扑空间上的爱因斯坦度量的存在性问题显然是几何中的最重要的问题之一.对应于紧致特殊和乐群的度量是爱因斯坦度量.除了凯勒度量,我们对这些度量的模空间知之甚少.对没有特殊结构要求的爱因斯坦度量,我们仅仅对四维流形知道一些拓扑限制.Berger[36], Gray [191] 与 Hitchin [230] 利用流形的欧拉数与符号差(signature) 建立了一些不等式.(这些工作当然是建立在陈先生[108]的用曲率形式表示示性类的工作之上)Gromov[201]用他的体积概念给出了更进一步的拓扑限制.LeBrun [287] 引进了 Seiberg-Witten 不变量的思想得到了更强的拓扑限制,在四维爱因斯坦流形上给出了漂亮的刚性定理.然而当爱因斯坦度量不具有特殊结构的时候,要理解他们的模空间就显得十分困难.例如,一个看上去十分简单的问题是在四维球面上是否只有一个爱因斯坦度量.M.Wang 与 Ziller[521] 利用对称约化在高维流形上给出了许多显式的爱因斯坦度量.爱因斯坦度量确实是几何分析中一个有趣的研究领域.

基本哲学:





由自然微分方程所支配的函数与子簇为几何结构提供了的深刻洞察.这些信息将会给出构造几何结构的方法.它也为物理学、代数几何学、拓扑学提供了重要的信息.反过来,从这些领域获取思想也是非常重要的.

除了这些基本原则外,还有用来理解空间扭曲的不变量,这是由陈类[108]提供的.陈类已应用在数学的各个分支及理论物理中.目前,我们差不多可以了解第一陈类的解析含义.但是几何学家要了解高阶陈形式的几何含义还有很长的路要走.陈类的解析表达式开辟了整体几何的新领域.陈省身教授对数学的影响是永恒的.

References

[1] ALLARD W K. On the first variation of a varifold: boundary behavior[J]. Ann. of Math. (2) 101 (1975), 418-446.

[2] ALMGREN J R F J. Almgren's big regularity parper. Q-valued functions minimizing Dirichlet's integral and the regularity of area-minimizing rectifiable currents up to codimension 2. With a preface by Jean E. Taylor and Vladimir Scheffer. World Scientific Monograph Series in Mathematics, 1[M]. River Edge: World Scientific Publishing Co., Inc., 2000.

[3] ALMGREN F, SIMON L. Existence of embedded solutions of Plateau's problem [J]. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 6 (1979), no. 3, 447-495.

[4] ANDERSON M T. The Dirichlet problem at infinity for manifolds of negative curvature[J]. J. Differential Geom. 18 (1983), no. 4, 701-721 (1984).

[5] ANDERSON M T, SCHOEN R. Positive harmonic functions on complete manifolds of negative curvature[J]. Ann. of Math. (2) 121 (1985), no. 3, 429-461.

[6] ATIYAH M. New invariants of 3-and 4-dimensional manifolds. The mathematical heritage of Hermann Weyl[M]. Providence: Amer. Math. Soc., 1988.

[7] ATIYAH M F, Bott R. The Yang-Mills equations over Riemann surfaces[J]. Philos. Trans. Roy. Soc. London Ser. A 308 (1983), no. 1505, 523-615.

[8] ATIYAH M F, Bott R. The moment map and equivariant cohomology[J]. Topology 23 (1984), no. 1, 1-28.

[9] ATIYAH M, Bott R, Patodi V K. On the heat equation and the index theorem[J]. Invent. Math. 19 (1973), 279-330.

[10] ATIYAH M F, Donnelly H, Singer I M. Eta invariants signature defects of cusps, and values of L-functions[J]. Ann. of Math. (2) 118 (1983), no. 1, 131-177.

[11] TIYAH M F, HITCHIN N J, SINGER I M. Self-duality in four-dimensional Riemanian geometry[J]. Proc. Roy. Soc. London Ser. A 362 (1978), no. 1711, 425-461.

[12] ATIYAH M F, PATODI V K, SINGER I M. Spectral asymmetry and Riemannian geometry. I[J]. Math. Proc. Cambridge Philos. Soc. 77 (1975), 43-69.

[13] ATIYAH M F, PATODI V K, SINGER I M. Spectral asymmetry and Riemannian



geometry. II[J]. Math. Proc. Cambridge Philos. Soc. 78(1975), no. 3, 405-432.

[14] ATIYAH M F, PATODI V K, SINGER I M. Spectral asymmetry and Riemannian geometry. III[J]. Math. Proc. Cambridge Philos. Soc. 79(1976), no. 1, 71-99.

[15] ATIYAH M F, SINGER I M. The index of elliptic operators on compact manifolds [J]. Bull. Amer. Math. Soc. 69(1963), 422-433.

[16] AUBIN T. Équations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire[J]. J. Math. Pures Appl. (9)55(1976), no. 3, 269-296.

[17] AUBIN T. Problèmes isopérimétriques et espaces de Sobolev[J]. J. Differential Geometry 11(1976), no. 4, 573-598.

[18] AUBIN T. Équations du type Monge-Ampère sur les variétés kähleriennes compactes[J]. C. R. Acad. Sci. Paris Sér. A-B 283(1976), no. 3, Aiii, A119-A121.

[19] AUBIN T. Espaces de Sobolev sur les variétés riemanniennes[J]. Bull. Sci. Math. (2)100(1976), no. 2, 149-173.

[20] BAILY W L Jr, BOREL A A. Compactification of arithmetic quotients of bounded symmetric domains[J]. Ann. of Math. (2)84(1966)442-528.

[21] BALLMANN W. On the Dirichlet problem at infinity for manifolds of nonpositive curvature[J]. Forum Math. 1(1989), no. 2, 201-213.

[22] BALLMANN W, LEDRAPPIER F. The Poisson boundary for rank one manifolds and their cocompact lattices[J]. Forum Math. 6(1994), no. 3, 301-313.

[23] BALLMANN W, SWIATKOWSIK J. On L^2-cohomology and property(T) for automorphism groups of polyhedral cell complexes[J]. Geom. Funct. Anal. 7(1997), no. 4, 615-645.

[24] BALLMANN W, THORBERGSSON G, ZILLER W. Closed geodesics on positively curved manifolds[J]. Ann. of Math. (2)116(1982), no. 2, 213-247.

[25] BANDO S. On the classification of three-dimensional compact Kaehler manifolds of nonnegative bisectional curvature[J]. J. Differential Geom. 19(1984), no. 2, 283-297.

[26] BANDO S, MABUCHI T. Uniqueness of Einstein Kähler metrics modulo connected group actions[M]. Amsterdam: Adv. Stud. Pure Math., 1987.

[27] BANGERT V, LONG Y M. The existense of two closed geodesics on every Finsler 2-sphere. Preprint(2005). Submitted.

[28] BARTNIK R. Existence of maximal surfaces in asymptotically flat spacetimes[J].

Comm. Math. Phys. 94(1984), no. 2,155-175.

[29] BAO D, Chern S S, SHEN Z. An introduction to Riemann-Finsler geometry. Graduate Texts in Mathematics, 200[M]. New York: Springer-Verlag,2000.

[30] BARTNIK R. Quasi-spherical metrics and prescribed scalar curvature[J]. J. Differential Geom. 37(1993), no. 1,31-71.

[31] BARTNIK R, MCKINNON J. Particlelike solutions of the Einstein-Yang-Mills equations[J]. Phys. Rev. Lett.61(1988), no. 2, 141-144.

[32] BARTOLOMEIS P, Tian G. Stability of complex vector bundles[J]. J. Differential Geom.43(1996), no. 2,231-275.

[33] BEAUVILLE A. Variétés Kähleriennes dont la première classe de Chern est nulle [J]. J. Differential Geom. 18(1983), no. 4, 755-782(1984).

[34] BECKER K, DASGUPTA K. Heterotic strings with torsion[J]. J. High Energy Phys. 11(2002),006,30 pp.

[35] BERGER M. Sur les groupes d'holonomie homogène des variétés à connexion affine et des variétés riemanniennes[J]. Bull. Soc. Math. France 83(1955), 279-330.

[36] BERGER M. Sur quelques variétés d'Einstein compactes[J]. Ann. Mat. Pura Appl.(4)53(1961),89-95.

[37] BERSHADSKY M, CECOTTI S, OOGURI H, et al. Kodaira-Spencer theory of gravity and exact results for quanturm string amplitudes[J]. Comm. Math. Phys. 165(1994), no.2,311-427.

[38] BESSON G, COURTOIS G, GALLOT S. Entropies et rigidités des espaces localement symétriques de courbure strictement négative[J]. Geom. Funct. Anal. 5 (1995), no.5,731-799.

[39] BISMUT J M. The Atiyah-Singer index theorem for families of Dirac operators: two heat equation proofs[J]. Invent. Math.83(1986), no. 1, 91-151.

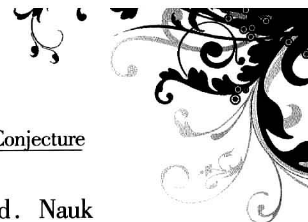
[40] BISMUT J M, GOETTE S. Equivariant de Rham torsions[J]. Ann. of Math.(2) 159(2004), no.1,53-216.

[41] Bismut J M, ZHANG W P. An extension of a theorem by Cheeger and Müller. With and appendix by Francois Laudenbach[J]. Astérisque No. 205(1992),235.

[42] BISMUT J M, ZHANG W P. Milnor and Ray-Singer metrics on the equivariant determinant of a flat vector bundle[J]. Geom. Funct. Anal.4(1994), no.2,136-212.

[43] BOCHNER S. Curvature and Betti numbers[J]. Ann. of Math.(2) 49(1948). 379-390.





[44] BOGOMOLOV F A. Hamiltonian Kälerian manifolds[J]. Dokl. Akad. Nauk SSSR 243(1978), no.5, 1101-1104.

[45] BOMBIERI E, GIUSTI E. Harnack's inequality for elliptic differential equations on minimal surfaces[J]. Invent. Math. 15(1972), 24-46.

[46] BOMBIERI E, De GIORGI E, GIUSTI E. Minimal cones and the Bernstein problem[J]. Invent. Math. 7(1969), 243-268.

[47] BORCHERDS R E. Automorphic forms on $O_{s+2,2}(R)$ and infinite products[J]. Invent. Math. 120(1995), no. 1, 161-213.

[48] BORCHERDS R E. The moduli space of Enriques surfaces and the fake Monster Lie superalgebra[J]. Topology 35(1996), no. 3, 699-710.

[49] BOTT R. A residue formula for holomorphic vector-fields[J]. J. Differential Geometry 1(1967), 311-330.

[50] BOTT R, TAUBES C H. Clifford on the rigidity theorems of Witten[J]. J. Amer. Math. Soc. 2(1989), no.1, 137-186.

[51] BOUTET DE MONVEL L, SJÖSTRAND I. Sur la singularité des noyaux de Bergman et de Szegö[J]. Soc. Math. France, Paris, 1976, 123-164.

[52] BOYER C P, GALICKI K, Kollár J. Einstein Metrics on Spheresmath[J]. arXiv: math. DG/0309408.

[53] BRAAM P J, DONALDSON S K. Floer's work on instanton homology, knots and surgery [J]. The Floer memorial volume, 195-256, Progr. Math. 133, Birkhäuser, Basel, 1995.

[54] BRAKKE K A. The motion of a surface by its mean curvatur. Mathematical Notes, 20[J]. Princeton University Press, Princeton, N.J., 1978.

[55] BRAY H L. Proof of the Riemannian Penrose inequality using the positive mass theorem[J]. J. Differential Geom. 59(2001), no. 2, 177-267.

[56] BROWN J D, YORK J W. Quasilocal energy and conserved charges derived from the gravitational action[J]. Phys. Rev. D(3)47(1993), no.4, 1407—1419.

[57] BRYANT R. Conformal and minimal immersions of compact surfaces into the 4-sphere[J]. J. Differential Geom. 17(1982), no.3, 455—473.

[58] BRYANT R L. Metrics with exceptional holonomy[J]. Ann. of Math. (2) 126 (1987), no. 3, 525-576.

[59] BRYANT R, SALAMON S. On the construction of some complete metrics with exceptional holonomy[J]. Duke Math. J. 58(1989), 829-850.

[60] BUCHDAHL N P. Hermitian-Einstein connections and stable vector bundles over compact complex surfaces[J]. Math. Ann. 280(1988), no. 4, 625-648.

[61] BURAGO Y, GROMOV M, PERELMAN G A D. Aleksandrov spaces with curvatures bounded below[J]. Uspekhi Mat. Nauk 47(1992), no. 2(284), 3-51, 222; translation in Russian Math. Surveys 47(1992), no. 2, 1-58.

[62] BURSTALL F E, FERUS D, PEDIT F, PINKALL U. Harmonic tori in symmetric spaces and commuting Hamiltonian systems on loop algebras[J]. Ann. of Math. (2) 138(1993), no. 1, 173-212.

[63] BUTSCHER A. Regularizing a singular special Lagrangian variety[J]. Comm. Anal. Geom. 12(2004), no. 4, 733-791.

[64] CAFFARELLI L, HARDT R, SIMON L. Minimal surfaces with isolated singularities[J]. Manuscripta Math. 48(1984), no. 1-3, 1-18.

[65] CAFFARELLI L, NIRENBERG, SPRUCK J. The Dirichlet problem for nonlinear second order elliptic equations, III[J]. Functions of the eigenvalues of the Hessian. Acta Math. 155(1985), 261-301.

[66] CALABI E. On Kähler manifolds with vanishing canonical class. Algebraic geometry and topology [M]. A symposium in honor of S. Lefschetz, pp. 78-89. Princeton University Press, Princeton, N. J., 1957.

[67] CALABI E. Improper affine hyperspheres of convex type and a generalization of a theorem by K. Jörgens[J]. Michigan Math. J. 5(1958) 105-126.

[68] CALABI E, VESENTINI. Sur les variétés complexes compactes localement symétriques[J]. Bull. Soc. Math. France 87(1959) 311-317.

[69] CANDELAS P, HOROWITZ G T, STROMINGER A and WITTEN E. Vacuum configurations for superstrings[J]. Nuclear Phys. B 258(1985), no. 1, 46-74.

[70] CANDELAS P. de la Ossa, X. C., P. S. Green and L. Parkes, A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory[J]. Nuclear Phys. B359(1991), no. 1, 21-74.

[71] CAO H D. Deformation of Kähler metrics to Kähler-Einstein metrics on compact Kähler manifolds[J]. Invent. Math. 81(1985), no. 2, 359-372.

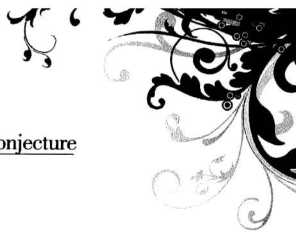
[72] CAO H D. On Harnack inequalities for the Kähler-Ricci flow[J]. Invent. Math. 109(1992), 247-263.

[73] CAO H D. Limits of solutions to the Kähler-Ricci flow[J]. J. Differential Geom. 45(1997), no. 2, 257-272.

[74] CAO J G, ESCOBAR J F. A new 3-dimensional curvature integral formula for PL-manifolds of non-positive curvature[J]. Comm. Anal. Geom. 11(2003), no. 3, 489-551.

[75] CAO J G, FAN H J, LEDRAPPIER F. Martin points on open manifolds of non-





positive curvature. To appear in Trans. AMS.

[76] CARDOSO G L, CURIO G G. Dall'Agata and D. Lüst, Heterotic string theory on non-Kähler manifolds with H-flux and gaugino condensate[J]. Proceedings of the 36th International Symposium Ahrenshoop on the Theory of Elementary Particles. Fortschr. Phys.52(2004),no.6-7,483-488.

[77] CATLIN D. The Bergman kernel and a theorem of Tian. Analysis and geometry in several complex variables(Katata,1997)[M]. Boston:Birkhäuser Boston, 1999.

[78] CHANG S C. Recent developments on the Calabi flow. Geometric evolution equations, 17-42,Contemp[J]. Math.,367, Amer. Math. Soc., Providence, RI, 2005.

[79] CHANG S Y A, GURSKY M J, YANG P. An a priori estimate for a fully nonlinear equation on four-manifolds. Dedicated to the memory of Thomas H. Wolff[J]. J. Anal. Math.87(2002),151-186.

[80] CHAU A, TAM L F. On the complex structure of Kähler manifolds with nonnegative curvature. ar Xiv:math. DG/0504422.

[81] CHEEGER J. A lower bound for the smallest eigenvalue of the Laplacian. Problems in analysis(Papers dedicated to Salomon Bochner, 1969)[M]. Princeton: Princeton Univ. Press, 1970.

[82] CHEEGER J. Analytic torsion and Reidemeister torsion[J]. Proc. Nat. Acad. Sci. U.S.A.74(1977),no.7,2651-2654.

[83] CHEEGER J, COLDING T H. On the structure of spaces with Ricci curvature bounded below. I[J]. J. Differential Geom.46(1997),no.3,406-480.

[84] CHEEGER J, COLDING T H. On the structure of spaces with Ricci curvature bounded below. II[J]. J. Differential Geom.54(2000),no.1,13-35

[85] CHEEGER J, COLDING T H. On the structure of spaces with Ricci curvature bounded below. III[J]. J. Differential Geom.54(2000),no.1,37-74.

[86] CHEEGER J, GROMOLL D. The splitting theorem for manifolds of nonnegative Ricci curvature[J]. J. Differential Geometry 6(1971/72),119-128.

[87] CHEN B L, TANG S H, ZHU X P. A uniformization theorem for complete noncompact Kähler surfaces with positive bisectional curvature[J]. J. Differential Geom. 67(2004),no.3,519-570.

[88] CHEN B L, ZHU X P. Complete Riemannian manifolds with pointwise pinched curvature[J]. Invent. Math.140(2000),no.2,423-452.

[89] CHEN B L, PING Z X. On complete noncompact Kähler manifolds with positive bisectional curvature[J]. Math. Ann. 327(2003),no.1,1-23.

[90] CHEN B L, ZHU X P. Volume growth and curvature decay of positive curved Kähler manifold, arXiv:math.DG/0211374.

[91] CHEN B L, ZHU X P. Flow with Surgery on Four-manifolds with Positive Isotropic Curvature. arXiv:math. DG/0504478.

[92] CHEN B L, ZHU X P. Uniqueness of the Ricci Flow on Complete Non compact Manifolds. arXiv:math. DG/0505447.

[93] CHEN C C, LIN C S. A spherical Harnack inequality for singualr solutions of nonlinear elliptic equations[J]. Ann. Sculoa Norm. Sup. Pisa Cl. Sci.(4)30 (2001),no.3-4,713-738(2002).

[94] CHEN M F, WANG F Y. Application of coupling method to the first eigenvalue on manifold[J]. Progr. Natur. Sci.(English Ed.)5(1995),no.2,227-229.

[95] CHEN X X. The space of Kähler metrics[J]. J. Differential Geom. 56(2000), no. 2,189-234.

[96] CHEN Y G, GIGA Y, GOTO S. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations[J]. J. Differential Geom.33(1991), no.3,749-786.

[97] CHEN S YG. Eigenvalue comparison theorems and its geometric applications[J]. Math. Z. 143(1975),no. 3,289-297.

[98] CHEN S YG. Eigenfunctions and nodal sets[J]. Comment. Math. Helv. 51 (1976),no.1,43-55.

[99] CHEN S YG, LI P, YAU S T. On the upper estimate of the heat kernel of a complete Riemannian manifold[J]. Amer. J. Math.103(1981),no.5,1021-1063.

[100] CHEN S YG, LI P, YAU S T. Heat equations on minimal submanifolds and their applications[J]. Amer. J. Math.106(1984),no.5,1033-1065.

[101] CHEN S YG, YAU S T. Differential equations on Riemannian manifolds and their geometric applications[J]. Comm. Pure Appl. Math.28(1975),no. 3, 333-354.

[102] CHEN S YG, YAU S T. Maximal space-like hypersurfaces in the Lorentz-Minkowski spaces[J]. Ann. of Math.(2)104(1976),no.3,407-419.

[103] CHEN S YG, YAU S T. On the regularity of the solution of the n-dimensional Minkowski problem[J]. Commun. Pure Appl. Math.29(1976),495-516.

[104] CHEN S YG, YAU S T. On the existence of a complete Kähler metric on non-compact complex manifolds and the regularity of Fefferman's equation [J]. Comm. Pure Appl. Math.33(1980),no.4,507-544.

[105] CHEN S YG, YAU S Y. The real Monge-Ampère equation and affine flat struc-



tures, in Proceedings of the 1980 Beijing Symposium on Differential Geometry and Differential Equations[J]. Science Press, Beijing, China(1982), Gordon and Breach Science Publishing Company, New York, pp. 339-370.

[106] CHEN S YG, YAU S T. Complete affine hypersurfaces. I. The completeness of affine metrics[J]. Comm. Pure Appl. Math.39(1986),no.6,839-866.

[107] Proceedings of the 1980 Beijing Symposium on Differential Geometry and Differential Equations. Vol. 1, 2, 3. Held in Beijing, August 18-September 21, 1980. Edited by S. S. Chern and Wen Tsün Wu[Wen Jun Wu]. Science Press, Beijing; Gordon and Breach Science Publishers, New York, 1982.

[108] CHERN S S. Characteristic classes of Hermitian manifolds[J]. Ann. of Math. (2)47(1946).85-121.

[109] CHERN S S, SIMONS J. Characteristic forms and geometric invariants[J]. Ann. of Math.(2)99(1974),48-69.

[110] CHOI H I, SCHOEN R. The space of minimal embeddings of a surface into a three-dimensional manifold of positive Ricci curvature[J]. Invent. Math. 81 (1985),no.3,387-394.

[111] CHOI H I, TREIBERGS A. Gauss maps of spacelike constant mean curvature hypersurfaces of Minkowski space[J]. J. Differential Geom. 32(1990),no.3, 775-817.

[112] CHOI H I, WANG A N. A first eigenvalue estimate for minimal hypersurfaces [J]. J. Differential Geom.18(1983),no. 3,559-562.

[113] CHOI S Y, GOLDMAN W M. The classification of real projective structures on compact surfaces[J]. Bull. Amer. Math. Soc. (N.S.)34(1997),no.2,161-171.

[114] CHOQUET-BRUHAT Y, MONCRIEF V. Future global in time Einsteinian spacetimes with U(1)isometry group[J]. Ann. Henri Poincare 2(2001),no.6, 1007-1064.

[115] CHOQUET-BRUHAT Y, YORK J W. On H. Friedrich's formulation of the Einstein equations with fluid sources[J]. Topol. Methods Nonlinear Anal. 18 (2001),no.2,321-335.

[116] CHRISTODOULOU D. The instability of naked singularities in the gravitational collapse of a scalar field[J]. Ann. of Math.(2)149(1999),no.1,183-217.

[117] CHRISTODOULOU D, KLAINERMAN S. The global nonlinear stability of the Minkowski space[M]. Princeton: Princeton Universtity Press, 1993.

[118] CHRISTODOULOU D, TAHVILDAR-ZADEH A. On the regularity of spherically

symmetric wave maps[J]. Comm. Pure Appl. Math. 46(1993), no. 7, 1041-1091.

[119] CHRISTODOULOU D, YAU S T. Some remarks on the quasi-local mass. Mathematics and general relativity, 9-14, Contemp[J]. Math., 71, Amer. Math. Soc., Providence, RI, 1988.

[120] CHRUUSCIEL P T. Semi-global existence and convergence of solutions of the Robinson-Trautman (2-dimensional Calabi) equation[J]. Comm. Math. Phys. 137(1991), no. 2, 289-313.

[121] CHUNG F. Discrete Isoperimetric Inequalities, Surveys in Differential Geometry, IX[M]. Inrernational Press, pp.53-82.

[122] CLEMENS C H. Double solids[J]. Adv. in Math.47(1983), no.2, 107-230.

[123] COLDING T H. Shape of manifolds with positive Rici curvature[J]. Invent. Math. 124(1996), no.1-3, 175-191.

[124] COLDING T H, MINICOZZI II W P. Harmonic functions on manifolds[J]. Ann. of Math.(2)146(1997), no. 3, 725-747.

[125] COLDING T H, MINICOZZI II W P. Embedded minimal surfaces without area bounds in 3-manifolds[M]. Providence: Amer. Math. Soc., 2000.

[126] COLDING T H, MINICOZZI II W P. The space of embedded minimal surfaces of fixed genus in a 3-manifold. I. Estimates off the axis for disks[J]. Ann. of Math(2)160(2004), no.1, 27-68.

[127] COLDING T H, MINICOZZI II W P. The space of embedded minimal surfaces of fixed genus in a 3-manifold. II. Multi-valued graphs in disks[J]. Ann. of Math.(2)160(2004), no.1, 69-92.

[128] COLDING T H, MINICOZZI II W P. The space of embedded minimal surfaces of fixed genus in a 3-manifold. III. Planar domains[J]. Ann. of Math.(2)160(2004), no.2, 523-572.

[129] COLDING T H, MINICOZZI II W P. The space of embedded minimal surfaces of fixed genus in a 3-manifold. IV. Locally simply connected[J]. Ann. of Math.(2)160(2004), no.2, 573-615.

[130] COLDING T H, MINICOZZI II W P. Estimates for the extinction time for the Ricci flow on certain 3-manifolds and a question of Perelman[J]. J. Amer. Math. Soc.18(2005), no.3, 561-569.

[131] COLDING T H, MINICOZZI II W P. An excursion into geomtric analysis[M]. Hong Kong: International Press, 83-146.

[132] VERDIÈRE COLIN DE. Y. Sur la multiplicitë de la première valeur propre non



nulle du laplacien[J]. Comment. Math. Helv. 61(1986), no.2, 254-270.

[133] CORLETTE K. Archimedean superrigidity and hyperbolic geometry[J]. Ann. of Math. (2)135(1992), no.1, 165-182.

[134] CROKE C B. Some isoperimetric inequalities and eigenvalue estimates[J]. Ann. Sci. ecole Norm. Sup. (4)13(1980), no.4, 419-435.

[135] CROKE C B. A sharp four-dimensional isoperimetric inequality[J]. Comment. Math. Helv. 59(1984), no. 187-192.

[136] DAFERMOS M. Stability and instability of the Cauchy horizon for the spherically symmetric Einstein-Maxwell-scalar field equations[J]. Ann. of Math. (2) 158 (2003), no. 3, 875-928.

[137] DAI X Z, WANG X D, WEI G F. On the stability of Riemannian manifold with parallel spinors[J]. Invent. Math. 161(2005), 151-176.

[138] DASKALOPOULOS G D, WENTWORTH R A. Geometric quantization for the moduli space of vector bundles with parabolic structure. Geometry, topology and physics[M]. Berlin: de Gruyter, 1997.

[139] DEMAILLY J. Holomorphic Morse inequalities. Several complex variables and complex geometry[M]. Providence: Amer. Math. Soc., 1991.

[140] DEMAILLY J. A numerical criterion for very ample line bundles[J]. J. Differential Geom. 37(1993), no.2, 323-374.

[141] DEMAILLY J. Effective bounds for very ample line bundles[J]. Invent. Math. 124(1996), no.1-3, 243-261.

[142] DIJKGRAAF R, VERLINDE H, VERLINDE E. Topological strings in d < 1[J]. Nuclear Phys. B 352(1991), no.1, 59-8.

[143] DONALDSON S K. An application of gauge theory to four-dimensional topology [J]. J. Differential Geom. 18(1983), no. 2, 279-315.

[144] DONALDSON S K. Anti self-dual Yang-Mills connections over complex algebraic surfaces and stable vector bundles[J]. Proc. London Math. Soc. (3)50(1985), no.1, 1-26.

[145] DONALDSON S K. Infinite determinants, stable bundles and curvature[J]. Duke Math. J. 54(1987), no.1, 231-247.

[146] DONALDSON S K. The Seiberg-Witten equations and 4-manifold topology[J]. Bull. Amer. Math. Soc. (N.S.)33(1996), no.1, 45-70.

[147] DONALDSON S K. Symplectic submanifolds and almost-complex geometry[J]. J. Differential Geom. 44(1996), no.4, 666-705.

[148] DONALDSON S K. Lefschetz pencils on symplectic manifolds[J]. J. Differential

Geom. 53(1999), no.2, 205-236.

[149] DONALDSON S K. Scalar curvature and projective embeddings. I[J]. J. Differential Geom. 59(2001), no.3, 479-522.

[150] DONALDSON S K. Scalar curvature and stability of toric varieties[J]. J. Differential Geom. 62(2002), no.2, 289-349.

[151] DONALDSON S K, FRIEDMAN R D. Connected Sums of Self-Dual Manifolds and Deformations of Singular Spaces[J]. Nonlinearity, 2(1989), 197-239.

[152] DONG R T. Nodal sets of eigenfunctions on Riemann surfaces[J]. J. Differential Geom. 36(1992), no.2, 493-506.

[153] DONNELLY H, FEFFERMAN C. Nodal sets of eigenfunctions on Riemannian manifolds[J]. Invent. Math. 93(1988), no. 1, 161-183.

[154] DUISTERMAAT J J, HÖRMANDER L. Fourier integral operators. II[J]. Acta Math. 128(1972), no.3-4, 183-269.

[155] DUISTERMAAT J J, GUILLEMIN V W. The spectrum of positive elliptic operators and periodic bicharacteristics[J]. Invent. Math. 29(1975), no.1, 39-79.

[156] EELLS JR J, SAMPSON J H. Harmonic mappings of Riemannian manifolds[J]. Amer. J. Math. 86(1964), 109-160.

[157] EIN L, LAZARSFELD R. Global generation of pluricanonical and adjoint linear series on smooth projective threefolds[J]. J. Amer. Math. Soc. 6(1993), no.4, 875-903.

[158] EVANS L C, SPRUCK J. Motion of level sets by mean curvature[J]. I. J. Differential Geom. 33(1991), no.3, 635-681.

[159] FEDERER H, FLEMING W H. Normal and integral currents[J]. Ann. of Math. (2)72(1960), 458-520.

[160] FEFFERMAN C. The Bergman kernel and biholomorphic mappings of pseudoconvex domains[J]. Invent. Math. 26(1974), 1-65.

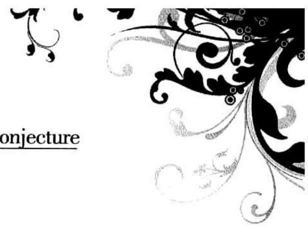
[161] FEFFERMAN C, PHONG D H. The uncertainty priinciple and sharp Garding inequalities[J]. Comm. Pure Appl. Math. 34(1981), no, 3, 285-331.

[162] FINSTER F, KAMRAN N, SMOLLER J, et al. Decay rates and probability estimates for massive Dirac particles in the Kerr-Newman black hole geometry[J]. Comm. Math. Phys. 230(2002), no.2, 201-244.

[163] FISCHER-COLBRIE D, SCHOEN R. The structure of complete stable minimal sufaces in 3-manifolds of nonnegative scalar curvature[J]. Comm. Pure Appl. Math. 33(1980), no.2, 199-211.

[164] FLOER A. An instanton-invariant for 3-manifolds[J]. Comm. Math. Phys. 118





(1988), no. 2, 215-240.

[165] FLOER A. Morse theory for Lagrangian intersections[J]. J. Differential Geom. 28(1988), no. 3, 513-547.

[166] FLOER A. Witten's complex and infinite-dimensional Morse theory[J]. J. Differential Geom. 30(1989), no. 1, 207-221.

[167] FREED D S. Characteristic numbers and generalized path integrals. Geometry, topology and physics, 126-138, Conf. Proc. Lecture Notes Geom. Topology, IV [M]. Cambridge: Internat. Press, 1995.

[168] FREEDMAN M. The topology of four-dimensional manifolds[J]. J. Differential Geom. 17(1982), no. 3, 357-453.

[169] FREEDMAN M, HASS J and SCOTT P. Least area incompressible surfaces in 3-manifolds[J]. Invent. Math. 71(1983), no. 3, 609-642.

[170] FREIR A. On the Martin boundary of Riemannian products[J]. J. Differential GEOM. 33(1991), no. 1, 215-232.

[171] FRIEDMAN R. Simultaneous resolution of threefold double points[J]. Math. Ann. 274(1986), no. 4, 671-689.

[172] FRIEDRICH H. Hyperbolic reductions for Einstein's equatons[J]. Classical Quantum Gravity. 13(1996), no. 6, 1451-1469.

[173] FU J X, YAU S T. Existence of Supersymmetric Hermitian Metrics with Torsion on Non-Kaehler Manifolds. arXiv: hep-th/0509028.

[174] FUJITA T. On polarized manifolds whose adjoint bundles are not semipositive [M]. Amsterdam: Pure Math, 1987.

[175] FUKAYA K. Morse homotopy, A^{∞} – category, and Floer homologies. Proceedings of GARC Workshop on Geometry and Topology'93(Seoul, 1993), 1 – 102, Lecture Notes Ser[M], Seoul: Seoul Nat. Univ., 1993.

[176] FUKAYA K, OH Y G. Zero-loop open strings in the cotangent bundle and Morse homotopy[J]. Asian J. Math. 1(1997), no. 1, 96-180.

[177] FUKAYA K, ONO K. Arnold conjecture and Gromov-Witten invariant[J]. Topology 38(1999), no. 5, 933-1048.

[178] FUTAKI A. An obstruction to the existence of Einstein? metrics[M]. Inv. Math., 73(1983), 437-443.

[179] GARLAND H. p-adic curvature and the ohomology of discrete subgroups of p-adic groups[M]. Ann. of Math. (2)97(1973), 375-423.

[180] GAUNTLETT J P, MARTELLI D, SPARKS J, et al. Sasaki-Einstein metrics on $S^2 \times S^3$[J]. Adv. Theor. Math. Phys. 8(2004), no. 4, 711-734.

[181] GERHARDT C. H-surfaces in Lorentzian manifolds[J]. Comm. Math. Phys. 89 (1983), no.4, 523-553.

[182] GERHARDT C. Hypersurfaces of prescribed scalar curvature in Lorentzian manifolds[J]. J. Reine Angew. Math.554(2003), 157-199.

[183] GEROCH R. Energy extraction[J]. Ann. New York Acad. Sci., 224, 108-117.

[184] GERVER M L and NADIRASHVILI N S. Condition A for isometry of Riemannian metrics in the disc[J]. Dokl. Akad. Nauk SSSR 275(1984), no.2, 289-293.

[185] GIESEKER D. On moduli of vector bundles on an algebraic surface[J], Ann. of Math. 106(1977)45-60.

[186] GIVENTAL A B. Equivariant Gromov-Witten invariants[J]. Internat. Math. Res. Notices 1996, no. 13, 613-663.

[187] GOLDMAN W M. On the polynomial cohomology of affine manifolds[J]. Invent. Math.65(1981/82), no.3, 453-457.

[188] GOMPF R E. Toward a topological characterization of symplectic manifolds[J]. J. Symplectic Geom.2(2004), no. 2, 177-206.

[189] GORDON C. WEBB D and WOLPERT S, Isospectral plane domains and surfaces via Riemannian orbifolds[J]. Invent. Math.110(1992), no.1, 1-22.

[190] GORESKY M, HARDER G, MACPHERSON R. Weighted cohomology[J]. Invent. Math.116(1994), no.1-3, 139-213.

[191] GRAY A. Invariants of curvature operators of four-dimensional Riemannian manifolds. Proceedings of the Thirteenth Biennial Seminar of the Canadian Mathematical Congress Vol. 2, pp.42-65. Canad. Math. Congr., Montreal, Que., 1972.

[192] GREENE R E, WU H. Analysis on nonocompact Kähler manifolds. Proc. Sympos. Pure Math., Vol. XXX, Part 2[M]. Providence: Amer. Math. Soc., 1977.

[193] GRIGOR'YAN A. Heat kernel of a nonocompact Riemannian manifold[M]. Providence: Amer. Math. Soc., 1995.

[194] GRIGOR'YAN A, NETRUSOV Y, YAN S T. Eigenvalues of elliptic operators and geometric applications, in Servey in Differential Geometry, V[M]. International Press.

[195] GRIGOR'YAN A, SALOFF-COSTE L. Dirichlet heat kernel in the exterior of a compact set[J]. Comm. Pure Appl. Math.55(2002), no.1, 93-133.

[196] GROMOLL D, MEYER W. On complete open manifolds of positive curvature [J]. Ann. of Math.(2)90(1969)75-90.



[197] GROMOLL D, MEYER W. Periodic geodesics on compact riemannian manifolds [J]. J. Differential Geometry 3(1969),493-510.

[198] GROMOLL D, WOLF J A. Some relations between the metric structure and the algebraic structure of the fundamental group in manifolds of nonpositive curvature [J]. Bull. Amer. Math. Soc.77(1971),545-552.

[199] GROMOV M. Manifolds of negative curvature[J]. J. Differential Geom. 13 (1978),no.2,223-230.

[200] GROMOV M C. Curvature, diameter and Betti numbers[J]. Comment. Math. Helv.56(1981),no.2,179-195.

[201] GROMOV M. Volume and bounded cohomology[J]. Inst. Hautes Études Sci. Publ. Math.56(1982),5-99.

[202] GROMOV M. Pseudoholomorphic curves in symplectic manifolds[J]. Invent. Math.82(1985),no.2,307-347.

[203] GROMOV M. Metric structures for Riemannian and non-Riemannian spaces Based on the 1981 French original[M]. Boston:Birkhäuser Boston, Inc.,1999.

[204] GROMOV M, LAWSON JR H B. Spin and scalar curvature in the presence of a fundamental group. I[J]. Ann. of Math.(2)111(1980),no 2,209-230.

[205] GROMOV M, LAWSON JR H B.The classification of simply connected manifolds of positive scalar curvature[J]. Ann. of Math.(2)111(1980),no.3,423-434.

[206] GROMOV M, SCHOEN R. Harmonic maps into singular spaces and p-adic superrigidity for lattices in groups of rank one[J]. Inst. Hautes Études Sci. Publ. Math.76(1992),165-246.

[207] Grove K, SHIOHAMA K. A generalized sphere theorem[J]. Ann. Math.(2) 106(1977),no.2,201-211.

[208] GU C H. On the initial-boundary value problem for harmonic maps from the 2-dimensional Minkowski space[J]. Manuscripta Math.33(1980/81),no.1,51-58.

[209] GUAN B, GUAN P. Hypersurfaces with prescribed curvatures[J]. Ann. Math. 156(2002),655-674.

[210] GUAN P F, MA X N. The Christoffel-Minkowski problem. I. Convexity of solutions of a Hessian equation[J]. Invent. Math.151(2003),no.3,553-577.

[211] Guillemin V. Wave-trace invariants[J]. Duke Math. J.83(1996),no.2,287-352.

[212] GUKOV S. Three-dimensional quantum gravity, Chern-Simons theory, and the A-polynomial[J]. Comm. Math. Phys. 255(2005),no.3,577-627.

[213] Gulliver R. Regularity of minimizing surfaces of prescribed mean curvature[J].

Ann. of Math. (2)97(1973), 275-305.

[214] HAKEN W. Various aspects of the three-dimensional Poincarè problem[J]Proc. Inst., Univ. of Georgia, Athens, Ga., 1969.

[215] HAMILTON R S. Three-manifolds with positive Ricci curvature[J]. J. Diferential Geom. 17(1982), no. 2, 255-306.

[216] HAMILTON R S. The Ricci flow on surfaces[M]. Providence: Amer. Math. Soc., 1988.

[217] HAMILTON R S. The Harnack estimate for the Ricci flow[J]. J. Differential Geom. 37(1993), no. 1, 225-243.

[218] HAMILTON R S. Four-manifolds with positive isotropic curvature[J]. Comm. Anal. Geom. 5(1997), no. 1, 1-92.

[219] HAN Q, HARDT R, LIN F H. Geometric measure of singular sets of elliptic equations[J]. Comm. Pure Appl. Math. 51(1998), no. 11-12, 1425-1443.

[220] Han Q, Hong J X, LIN C S. Local isometric embedding of surfaces with nonpositive Gaussian curvature[J]. J. Differential Geom. 63(2003), no. 3, 475-520.

[221] HARDT R, Simon L. Boundary regularity and embedded solutions for the oriented Plateau problem[J]. Ann. of Math. (2)110(1979), no. 3, 439-486.

[222] HARVEY F R, LAWSON JR H B. On boundaries of complex analytic varieties. I[J]. Ann. of Math. (2)102(1975), no. 2, 223-290.

[223] HARVEY F R, LAWSON JR H B. On boundaries of complex analytic varieties. II[J]. Ann. Math. (2)106(1977), no. 2, 213-238.

[224] HARVEY R, LAWSON JR H B. Calibrated geometries[J]. Acta Math. 148 (1982), 47-157.

[225] HASSELL A, MELROSE R, VASY A. Spectral and scattering theory for symbolic potentials of order zero[J]. Adv. Math. 181(2004), no. 1, 1-87.

[226] HERSCH J. Quatre propriétés isopérimétriques de membranes sphériques homogènes[J]. C. R. Acad. Sci. Paris Sér. A-B 270(1970), A1645-A1648.

[227] HINGSTON N. Equivariant Morse theory and closed geodesics[J]. J. Differential Geom. 19(1984), no. 1, 85-116.

[228] HIRZEBRUCH F, KODAIRA K. On the complex projective spaces[J]. J. Math. Pures Appl. (9)36(1957), 201-216.

[229] HITCHIN H. Harmonic spinors[J]. Advances in Math. 14(1974), 1-55.

[230] HITCHIN N. Compact four-dimensional Einstein manifolds[J]. J. Differential Geometry 9(1974), 435-441.

[231] HITCHIN N J. The self-duality equations on a Riemann surface[J]. Proc. Lon-



don Math. Soc. (3)5(1987), no. 1, 59-126.

[232] HITCHIN N J. Harmonic maps from a 2-torus to the 3-sphere[J]. J. Differential Geom. 31(1990), no. 3, 627-710.

[233] HITCHIN N J. Lie groups and Teichmuller space[J]. Topology 31(1992), no. 3, 449-473.

[234] HITCHIN N J. The moduli space of special Lagrangian submanifolds[J]. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)25(1997), no. 3-4, 503-515.

[235] HITCHIN N. The geometry of three-forms in six dimensions[J]. J. Differential Geom. 55(2000), no. 3, 547-576.

[236] HITCHIN N. Stable forms and special metrics. Global differential geometry: the mathematical legacy of Alfred Gray (Bilbao, 2000), 70-89, Contemp [J]. Math., 288, Amer. Math. Soc., Providence, RI, 2001.

[237] HOFER H, WYSOCKI K, ZEHNDER E. The dynamics on three-dimensional strictly convex energy surfaces[J]. Ann. of Math. (2) 148(1998), no. 1, 197-289.

[238] HOFFMAN-OSTENHOF M, HOFFMANN-OSTENHOF T, NADIRASHVILI N. Critical sets of smooth solutions to elliptic equations in dimension 3, Indiana Univ[J]. Math. J. 45(1996), 15-37.

[239] HONG J X. Realization in R^3 of complete Riemannian manifolds with negative curvature[J]. Comm. Anal. Geom. 1(1993), no. 3-4, 487-514.

[240] HOPKINS M J. Topological modular forms, the Witten genus, and theorem of the cube. Proceedings of the International Congress of Mathematicians, Vol. 1, 2 [M]. Basel: Birkhäuser Basel, 1995.

[241] HÖRMANDER L. An introduction to complex analysis in several variables[M]. Third edition. Amsterdam: North-Holland Publishing Co., 1990.

[242] HUA L K. Harmonic analysis of functions of several variables in the classical domains[J]. Providence: American Mathematical Society, 1963.

[243] UISKEN G. Flow by mean curvature of convex surfaces into spheres[J]. J. Differential Geom. 20(1984), no. 1, 237-266.

[244] UISKEN G. Contracting convex hypersurfaces in Riemannian manifolds by their mean curvature[J]. Invent. Math. 84(1986), no. 3, 463-480.

[245] UISKEN G, ILMANEN T. The inverse mean curvature flow and the Riemannian Penrose inequality[J]. J. Differential Geom. 59(2001), no. 3, 353-437.

[246] UISKEN G, SINESTRARI C. Mean curvature flow singularities for mean convex surfaces[J]. Calc. Var. Partial Differential Equations 8(1999), no. 1, 1-14.

[247] UISKEN G, SINESTRARI C. Convexity estimates for mean curvature flow and singularities of mean convex surfaces[J]. Acta Math. 183(1999), no. 1, 45-70.

[248] UISKEN G, YAN S T. Definition of center of mass for isolated physical systems and unique foliations by stable spheres with constant mean curvature[J]. Invent. Math. 124(1996), no. 1-3, 281-311.

[249] Iwaniec H, SARNAK P. L^{∞} norms of eigenfunctions of arithmetic surfaces[J]. Ann. of Math. (2) 141(1995), no. 2, 301-320.

[250] JI L Z, MACPHERSON R. Geometry of compactifications of locally symmetric spaces[J]. Ann. Inst. Fourier(Grenoble) 52(2002), no. 2, 457-559.

[251] JI M, WANG G Y. Minimal surfaces in Riemannian manifolds[J]. Mem. Amer. Math. Soc. 104(1993), no. 495.

[252] JÖRGENS K. Über die Losungen der Differentialgleichung $rt - s^2 = 1$[J]. Math. Ann. 127(1954), 130-134.

[253] JORGENSON J, TODOROV A. Enriques surfaces, analytic discriminants, and Borcherds's Φ function[J]. Comm. Math. Phys. 191(1998), no. 2, 249-264.

[254] JOST J. Embedded minimal surfaces in manifolds diffeomorphic to the threedimensional ball or sphere[J]. J. Differential Geom. 30(1989), no. 2, 555-577.

[255] JOST J. Convex functionals and generalized harmonic maps into spaces of nonpositive curvature[J]. Comment. Math. Helv. 70(1995), no. 4, 659-673.

[256] JOST J, YAU S T. Applications of quasilinear PDE to algebraic geometry and arithmatic lattices. Algebraic geometry and related topics, 169-193.

[257] JOST J, Yan S T. Harmonic maps and superrigidity. Differential geometry: partial differential equations on manifolds, 245-280, Proc. Sympos[J]. Pure Math., 54, Part 1, Amer. Math. Soc, 1993.

[258] JOYCE D D. Compact 8-manifolds with holonomy Spin(7)[J]. Invent. Math. 123(1996), no. 3, 507-552.

[259] JOYCE O D. Compact Riemannian 7-manifolds with holonomy G_2. I, II[J]. J. Differential Geom. 43(1996), no. 2, 291-328, 329-375.

[260] JOYCE D. A new construction of compact 8-manifolds with holonomy Spin(7)[J]. J. Differential Geom. 53(1999), no. 1, 89-130.

[261] KATOK A B. Ergodic properties of degenerate integrable Hamiltonian systems. Izv. Akad. Nauk SSSR. 37(1973)(Russian)[J]. Math. USSR-Izv. 7(1973) 535-571.

[262] KIM J, LEBRUN C, PONTECORVO M. Sclar-flat Kähler surfaces of all genera[J]. J. Reine Angew. Math. 486(1997), 69-95.





[263] KING J R. The currents defined by analytic varieties[J]. Acta Math. 127 (1971), no.3-4, 185-220.

[264] KLAINERMAN S, NICOLÒ. Peeling properties of asymptotically flat solutions to the Einstein vacuum equations[J]. Classical Quantum Gravity 20(2003), no. 14, 3215-3257.

[265] KLAINERMAN S, SELBERG S. Bilinear estimates and applications to nonlinear wave equations[J]. Commun. Contemp. Math.4(2002), no.2, 223-295.

[266] KLINGENBERG W. Lectures on closed geodesics[M]. Third edition. Bonn: Mathematisches Institut der Universität Bonn, 1977.

[267] KOBAYASHI S. On compact Kähler manifolds with positive definite Ricci tensor [J]. Ann. of Math.(2)74(1961)570-574.

[268] KODAIRA K. On a differential-geometric method in the theory of analytic stacks [J]. Proc. Nat. Acad. Sci. U. S. A.39(1953).1268-1273.

[269] KODAIRA K. On Kähler varieties of restricted type(an intrinsic characterization of algebraic varieties)[J]. Ann. of Math.(2)60(1954).28-48.

[270] KOLLÁR J. Effective base point freeness[J]. Math. Ann.296(1993), no. 4, 595-605.

[271] KONTSEVICH M. Enumeration of rational curves via torus actions[M]. Boston: Birkhäuser Boston, 1995.

[272] KONTSEVICH M. Feynman diagrams and low-dimensional topology[M]. Basel: Birkhäuser Basel, 1994.

[273] KONTSEVICH M. Homological algebra of mirror symmetry [M]. Basel: Birkhäuser Basel, 1995.

[274] KOREVAAR N. An easy proof of the interior gradient bound for solutions to the prescribed mean curvature equation[M]. Providence: Amer. Math. Soc., 1986.

[275] KOREVAAR N. Upper bounds for eigenvalues of conformal metrics[J]. J. Differential Geom.37(1993), no.1, 73-93.

[276] KOREVAAR N, MAZZEO R, PACARD F, et al. Refined asymptotics for constant scalar curvature metrics with isolated singularities[J]. Invent. Math. 135 (1999), no.2, 233-272.

[277] KOREVAAR N, SCHOEN R. Sobolev spaces and harmonic maps for metric space targets[J]. Comm. Anal. Geom.1(1993), no.3-4, 561-659.

[278] KRAMER P. Group actions on compact hyperbolic manifolds and closed geodesics[M]. Providence: Amer. Math. Soc., 2004.

[279] KRONHEIMER P B, MROWKA T S. Embedded surfaces and the structure of

Donaldson's polynomial invariants[J]. J. Differential Geom. 41(1995), no. 3, 573-734.

[280] KURANISHI M. Strongly pseudoconvex CR structures over small balls. I. An a priori estimate[J]. Ann. of Math. (2)115(1982), no. 3, 451-500.

[281] KURANISHI M. Strongly pseudoconvex CR structures over small balls. II. An regularity theorem[J]. Ann. of Math. (2)116(1982), no. 1, 1-64.

[282] KURANISHI M. Strongly pseudoconvex CR structures over small balls. III. An embedding theorem[J]. Ann. of Math. (2)116(1982), no. 2, 249-330.

[283] LAWSON H B, SIMONS J. On stable currents and their application to global problems in real and complex geometry[J]. Ann. of Math. (2)98(1973), 427-450.

[284] LAWSON JR H B, YAU S T. Compact manifolds of nonpositive curvature[J]. J. Differential Geometry 7(1972), 211-228.

[285] LAWSON JR H B, YAU S T. Scalar curvature, non-abelian group actions, and the degree of symmetry of exotic spheres[J]. Comment. Math. Helv. 49(1974), 232-244.

[286] LEBRUN C R. Explicit Self-Dual Metrics on $CP_2 \neq \cdots \neq CP_2$[J]. J. Diff. Geom. 34(1991), 223-253.

[287] LEBRUN C. Ricci curvature, minimal volumes, and Seiberg-Witten theory[J]. Invent. Math. 145(2001), no. 2, 279-316.

[288] LEE Y Y. Embedded special Lagrangian submanifolds in Calabi-Yau manifolds [J]. Comm. Anal. Geom. 11(2003), no. 3, 391-423.

[289] LEUNG N C C. Einstein type metrics and stability on vector bundles[J]. J. Differential Geom. 45(1997), no. 3, 514-546.

[290] LEUNG N C C. Lagrangian submanifolds in hyperKähler manifolds, Legendre transofrmation[J]. J. Differential Geom. 61(2002), no. 1, 107-145.

[291] LEUNG N C C. Topological quantum field theory for Calabi-Yau threefolds and G_2-manifolds[J]. Adv. Theor. Math. Phys. 6(2002), no. 3, 575-591.

[292] LI J. Algebraic geometric interpretation of Donaldson's polynomial invariants [J]. J. Differential Geom. 37(1993), no. 2, 417-466.

[293] LI J. Stable morphisms to singular schemes and relative stable morphisms[J]. J. Differential Geom. 57(2001), no. 3, 509-578.

[294] LI J. A degeneration formula of GW-invariants[J]. J. Differential Geom. 60(2002), no. 2, 199-293.

[295] LI J, LIU C C M, LIU K F, et al. A Mathematical Theory of the Topological



Vertex Authors, arXiv: math. AG/0408426.

[296] LI J, TIAN G. Virtual moduli cycles and Gromov-Witten invariants of algebraic varieties[J]. J. Amer. Math. Soc. 11(1998), no. 1, 119-174.

[297] LI J, TIAN G. Virtual moduli cycles and Gromov-Witten invariants of general symplectic manifolds[M]. Cambridge: Internat. Press, 1998.

[298] LI J, YAU S T. Hermitian-Yang-Mills connection on non-Kähler manifolds[M]. Singapore: World Sci Publishing, 1987.

[299] LI J, YAU S T. The Existence of Supersymmetric String Theory with Torsion. ar Xiv: hep-th/0411136.

[300] LI J, YAU S T, ZHENG F Y. A simple proof of Bogomolov's theorem on class VII_0 surfaces with $b_2=0$[J]. Illinois J. Math. 34(1990), no. 2, 217-220.

[301] LI P. A lower bound for the first eigenvalue of the Laplacian on a compact manifold[J]. Indiana Univ. Math. J. 28(1979), no. 6, 1013-1019.

[302] LI P. A lower bound for the first eigenvalue of the Laplacian on a compact manifold[J]. Indiana Univ. Math. Res. Lett. 4(1979), no. 1, 35-44.

[303] LI P, SCHOEN R, YAN S T. On the isoperimetric inequality for minimal surfaces[J]. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 11(1984), no. 2, 237-244.

[304] LI P, TIAN Q. On the heat kernel of he Bergmann metric on algebraic varieties [J]. J. Amer. Math. Soc. 8(1995), no. 4, 857-877.

[305] LI P, WANG J P. Counting dimensions of L-harmonic functions[J]. Ann. of Math. (2) 152(2000), no. 2, 645-658.

[306] LI P, WANG J P. Counting manifolds with positive spectrum. II[J]. J. Differential Geom. 62(2002), no. 1, 143-162.

[307] LI P, YAU S T. Estimates of eigenvalues of compact Riemannian manifold[M]. Providence: Amer. Math. Soc., 1980.

[308] LI P, YAU S T. A new conformal invariant and its application to the Willmore conjecture and the first eigenvalue of compact surfaces[J]. Invent. Math. 69 (1982), no. 2, 269-291.

[309] LI P, YAU S T. On the parabolic kernel of the Schrödinger operator[J]. Acta Math. 156(1986), no. 3-4, 153-201.

[310] LI P, YAU S T. Curvature and holomorphic mappings of complete Kähler manifolds[J]. Compositio Math. 73(1990), no. 2, 125-144.

[311] LI T J. Symplectic 4-manifolds with Kodaira dimension zero. Preprint.

[312] LI T J, LIU A. General wall crossing formula[J]. Math. Res. Lett. 2(1995),

no.6,797-810.

[313] LI T J, LIU A. Symplectic structure on ruled surfaces and a generalized adjuntion formula[J]. Math. Res. Lett.2(1995),no.4,453-471.

[314] LIAN B H, LIU K F, YAN S T. Miror principle. I[J]. Asian J. Math.1 (1997), no.4,729-763.

[315] LIAN B H, LIU K F, YAN S T. Mirror principle. II.Sir Michael Atiyah:a great mathematician of the twentieth century[J]. Asian J. Math.3(1999),no.1,109-146.

[316] LIAN B H, LIU K F, YAN S T. Mirror principle. III[J]. Asian J. Math. 3 (1999),no.4,771-800.

[317] LIAN B H, LIU K F, YAN S T. Mirror principle. IV[M]. Somerville:Int, Press, 2000.

[318] LICHNEROWICZ A. Spineurs harmoniques[J]. C. R. Acad. Sci. Paris, 257 (1963),7-9.

[319] LICHNEROWICZ A. Magnetohydrodynamics: waves and shock waves in curved space-time[M]. Dordrecht:Kluwer Academic Publishers Group, 1994.

[320] LIN C S. The local isometric embedding in $\mathbf{R}^3$ of 2-dimensional Riemannian manifolds with nonnegative curvature[J]. J. Differential Geometry 21(1985), 213-230.

[321] LIN C S. The local isometric embedding in R^3 of two-dimensional Riemannian manifolds with Gaussian curvature changing sign cleanly[J]. Comm. Pure Appl. Math.39(1986)867-887.

[322] LIN F H. Asymptotic behavior of area-minimizing currents in hyperbolic space [J]. Comm. Pure Appl. Math.42(1989),no.3,229-242.

[323] LIN F H. Nodal sets of solutions of elliptic and parabolic equatons[J]. Comm. Pure Appl. Math.44(1991),no.3,287-308.

[324] LIN F H. Gradient estimates and blow-up analysis for stationary harmonic maps [J]. Ann. of Math.(2)149(1999),no.3,785-829.

[325] LIN F H. Mapping problems, fundamental groups and defect measures[J]. Acta Math. Sin. (Engl. Ser.)15(1999),no.1,25-52.

[326] LINDBLAD H, RODNIANSKI I. Global existence for the Einstein vacuum equations in wave coordinates[J]. Comm. Math. Phys.256(2005),no.1,43-110.

[327] LIU A K. Some new applications of general wall crossing formula, Gompf's conjecture and its applications. Math. Res. Lett.3(1996),no.5,569-585.

[328] LIU C C M, YAU S T. Positivity of Quasilocal Mass[J]. Phys. Rev. Lett.90





(2003).

[329] LIU C C M, YAU S T. Positivity of quasi-local mass II. arXiv: math. DG/0412292.

[330] LIU G, TIAN G. Floer homology and Arnold conjecture[J]. J. Differential Geom.49(1998), no.1, 1-74.

[331] LIU K F. On modular invariance and rigidity theorems[J]. J. Differential Geom.41(1995), no.2, 343-396.

[332] LIU K F. On elliptic genera and theta-functions[J]. Topology 35(1996), no.3, 617-640.

[333] LIU K F, SUN X F, YAU S T. Ganonical metrics on the moduli space of Riemann surfaces. I[J]. J. Differential Geom.68(2004), no.3, 571-637.

[334] LIU K F, SUN X F, YAU S T. Ganonical Metrics on the Moduli Space of Riemann Surfaces. II. arXiv. math. DG/0409220.

[335] LOFTIN J C. Affine spheres and convex RP^n-manifolds[J]. Amer. J. Math.123 (2001), no.2, 255-274.

[336] LOFTIN J, YAU S T, ZASLOW E. Affine Manifolds, SYZ Geometry, and the “Y” Vertex Authors. arXiv: math. DG/0405061.

[337] LOOIJENGA E. L^2-cohomology of locally symmetric varieties[J]. Compositio Math. 67(1988), no.1, 3-20.

[338] LU Z Q. On the lower order terms of the asymptotic expansion of Tian-Yau-Zelditch[J]. Amer. J. Math.122(2000), 235-273.

[339] LU P, TIAN G. The complex structure on a connected sum of $S^3 \times S^3$ with trivial canonical bundle[J]. Math. Ann.298(1994), no.4, 761-764.

[340] Lubotzky A, PHILLIPS R, SARNAK P. Ramanujan graphs[J]. Combinatorica8 (1988), no.3, 261-277.

[341] LUO H. Geometric criterion for Gieseker-Mumford stabhility of polarized manifolds[J]. J. Differential Geom.49(1998), no.3, 577-599.

[342] LUSZTIG G. Novikov's higher signature and families of elliptic operators[J]. J. Differential Geometry 7(1972), 229-256.

[343] LUTWAK E. Selected affine isoperimetric inequalities. Handbook of convex geometry[M]. Amsterdam: North-Holland, 1993.

[344] LUTWAK E, YANG D and ZHANG G Y. L_p affine isoperimetric inequalities [J]. J. Differential Geom.56(2000), no.1, 111-132.

[345] MABUCHI T. K-energy maps ingegrating Futaki invariants[J]. Tohoku Math. J. (2)38(1986), no.4, 575-59.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

[346] MARGULIS G A. Discrete groups of motions of manifolds of nonpositive curvature[J]. Am. Math. Soc. 109(1977),33-45.

[347] MARGULIS G A. Explicit constructions of graphs without short cycles and low density codes[J]. Combinatorica 2(1982),no.1,71-78.

[348] MATSUSAKA I. Polarized varieties with a given Hilbert polynomial[J]. Amer. J. Math.94(1972),1027-1077.

[349] Matsushima Y. Sur la structure du group d'homeomorphismes analytiques d'une certaine varietie Kaehleriennes[J]. Nagoya Math. J.,11(1957),145-150.

[350] MATSUSHIMA Y. On the first Betti number of compact quotient spaces of higherdimensional symmetric spaces[J]. Ann. of Math.(2)75(1962),312-330.

[351] MCDUFF D. The structure of rational and ruled symplectic 4-manifolds[J]. J. Amer. Math.Soc.3(1990),no.3,679-712.

[352] MCKEAN H P JR, SINGER I M. Curvature and the eigenvalues of the Laplacian [J]. J. Differential Geometry 1(1967),no.1,43-69.

[353] MEEKS W H III, ROSENBERG H. The uniqueness of the helicoid[J]. Ann. of Math.(2)161(2005),no.2,727-758.

[354] MEEKS W H, SIMON L, YAU S T. Embedded minimal surfaces, exotic spheres, and manifolds with positive Ricci curvature[J]. Ann. of Math.(2)116 (1982),no.3,621-659.

[355] MEEKS W H, YAU S T. Topology of three-dimensional manifolds and the embedding problems in minimal surface theory[J]. Ann. of Math.(2)112(1980), no.3,441-484.

[356] MEEKS W H, YAU S T. The equivariant Dehn's lemma and loop theorem[J]. Comment. Math. Helv.56(1981),no.2,225-239.

[357] MEEKS W H, YAU S T. The existence of embedded minimal surfaces and the problem of uniqueness[J]. Math. Z.179(1982),no.2,151-168.

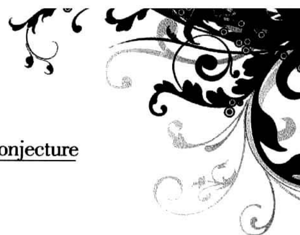
[358] MEEKS W H, YAU S T. The classical Plateau problem and the topology of three-dimensional manifolds. The embedding of the solution given by Douglas-Morrey and an analytic proof of Dehn's lemma[J]. Topology 21(1982),no.4, 409-442.

[359] MELAS A D. On the nodal line of the second eigenfunction of the Laplacian in R^2[J]. J. Differential Geom.35(1992),no.1,255-263.

[360] Melrose R. The inverse spectral problem for planar domains[M]. Canberra:Austral. Nat. Univ., 1996.

[361] MICALLEF M J. Stable minimal surfaces in Euclidean space[J]. J. Differential





Geom. 19(1984), no.1, 57-84.

[362] MICALLEF M J, MOORE J D. Minimal two-spheres and the topology of manifolds with positive curvature on totally isotropic two-planes[J]. Ann. of Math. (2)127(1988), no.1, 199-227.

[363] MICHELSOHN ML. Kähler manifolds with vanishing first Chern class. Seminar on Differential Geometry[J], Ann. of Math. Stud, 102, Princeton Univ. Press, Princeton, N.J., 1982.

[364] MICHELSOHN M L. On the existence of special metrics in complex geometry [J]. Acta Math. 149(1982), no.3-4, 261-295.

[365] MJYAOKA Y. On the Chern numbers of surfaces of general type[J]. Invent. Math. 42(1977), 225-237.

[366] MOK N. Courbure bisectionnelle positive et variétés algébriques affines. C R Acad[J]. Sci. Paris Sér. I Math. 296(1983), no.11, 473-476.

[367] MOK N. An embedding theorem of complete Kähler manifolds of positive bisectional curvature onto affine algebraic varieties[J]. Bull. Soc. Mat. France 112 (1984), 197-250.

[368] MOK N. The uniformization theorem for compact Kähler manifolds of nonnegative holomorphic bisectional curvature[J]. J Differential Geom. 27(1988), no.2, 179-214.

[369] MOK N, SIU Y T, YAU S T. The poincaré-Lelong equation on complete Kähler manifolds[J]. Compositio Math. 44(1981), no.1-3, 183-218.

[370] MOK N, SIU Y T, YEUNG S K. Geometric superrigidity[J]. Invent. Math. 113 (1993), no.1, 57-83.

[371] MOK N, YAU S T. Completeness of the Kähler-Einstein metric on bounded domains and the characterization of domains of holomorphy by curvature conditions [M]. Providence: Amer. Math. Soc., 1983.

[372] MOK N, ZHONG J Q. Curvature characterization of compact Hermitian symmetric spaces[J]. J. Differential Geom. 23(1986), no.1, 15-67.

[373] MOK N, ZHONG J Q. Compactifying complete Kähler-Einstein manifolds of finite topological type and bounded curvature[J]. Ann. of Math. (2)129(1989), no.3, 427-470.

[374] MORI S. Projective manifolds with ample tangent bundles[J]. Ann. of Math. (2)110(1979), no.3, 593-606.

[375] Müller W. Analytic torsion and R-torsion of Riemannian manifolds[J]. Adv. in Math. 28(1978), no.3, 233-305.

[376] Mostow G D. Strong rigidity of locally symmetric spaces[M]. Tokyo: University of Tokyo Press, 1973.

[377] MCMULLEN C T. The moduli space of Riemann surfaces is K··· ahler hyperbolic [J]. Ann. of Math.(2)151(2000), no.1, 327-357.

[378] Nadel A M. Multiplier ideal sheaves and Kähler-Einstein metrics of positive scalar curvature[J]. Ann. of Math.(2)132(1990), no.3, 549-596.

[379] NADIRASHVILI N. Hadamard's and Calabi-Yau's conjectures on negatively curved and minimal surfaces[J]. Invent. Math.126(1996), no.3, 457-465.

[380] NEEL R, STROOCK D. Analysis of the cutlocus via the heat kernel. Surveys in differential Geomtry IX[M]. Hong Kong: International Press, 337-350.

[381] NI L. Ancient solution to Kahler-Ricci flow. arXiv: math. DG/0502494.

[382] NIRENBERG L. The Weyl and Minkowski problems in differential geometry in the large[J]. Commun. Pure Appl. Math.6(1953), 337-394.

[383] NOVIKOV S P. Multivalued functions and functionals. An analogue of Morse theory[J]. Dokl. Akad. Nauk SSSR 260(1981), no.1, 31-35.

[384] OH Y G. Floer cohomology of Lagrangian intersections and pseudoholomorphic disks. III[M]. Basel: Birkhäuser 1995.

[385] OSHER S. A level set formulation for the solution of the Dirichlet problem for Hamilton-Jacobi equations[J]. SIAM J. Math. Anal. 24(1993), no. 5, 1145-1152.

[386] OSSERMAN R. A proof of the regularity everywhere of the classical solution to Plateau's problem[J]. Ann. of Math.(2)91(1970), 550-569.

[387] OZSVÁTH P, SZABOZ. The symplectic Thom conjecture[J]. Ann. of Math. (2)151(2000), no.1, 93-124.

[388] PARKER T H, WOLFSON J G. Pseudo-holomorphic maps and bubble trees (English. English summary)[J]J. Geom. Anal.3(1993), no.1, 63-98.

[389] PERELMAN G. The entropy formula for the Ricci flow and its geometric applications. arXiv: math. DG/0211159.

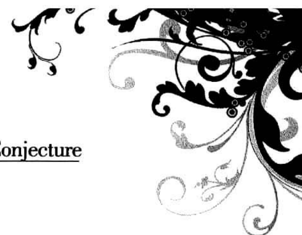
[390] PERELMAN G. RICCI flow with surgery on three-manifolds, arXiv: math. DG/0303109.

[391] PERELMAN G. inite extinction time for the solutions to the Ricci folw on certain three-manifolds, arXiv: math. DG/0307245.

[392] PESTO L, UHLMANN G. The boundary distance function and the Dirichlet-to-Neumann map[J]. Math. Res. Lett.11(2004), no.2-3, 285-297.

[393] PHONG D H, STURM J. On stability and the convergence of the Kähler-Ricci





flow. arXiv: math. DG/0412185.

[394] PHONG D H, STURM J. The Monge-Ampère operator and geodesics in the space of Kähler potentials. arXiv: math. DG/0504157.

[395] PITTS J T. Existence and regularity of minimal surfaces on Riemannian manifolds [M]. Tokyo: University of Tokyo Press, 1981.

[396] PITTS J T, RUBINSTEIN J H. Equivariant minimax and minimal surfaces in geometric three-manifolds[J]. Bull. Amer. Math. Soc. (N.S.) 19(1988), no. 1, 303-309.

[397] POGORELOV A V. On the question of the existence of a convex surface with a given sum principal radii of curvature[J]. Usp. Mat. Nauk 8(1953), 127-130.

[398] POGORELOV A V. On the improper convex affine hyperspheres[J]. Geometriae Dedicata 1(1972), no. 1, 33-46.

[399] POGORELOV A V. Existence of a closed convex hypersurface with prescribed curvature[J]. Dokl. Akad. Nauk SSSR 274(1984), no. 1, 28-31.

[400] POLYA G, XZEGÖ. Isoperimetric Inequalities in Mathematical Physics [M]. Princeton: Princeton University Press, 1951.

[401] QING J, TIAN G. On the uniqueness of the foliation of spheres of constant mean curvature in asymptotically flat 3-manifolds. arXiv: math. DG/0506005.

[402] QUILLEN D. Superconnections and the Chern character [J]. Topology 24 (1985), no. 1, 89-95.

[403] QUILLEN D. Cyclic cohomology and algebra extensions[J]. K-Theory 3(1989), no. 3, 205-246.

[404] ROSENBERG J. C^*-algebras, positive scalar curvature and the Novikov conjecture[J]. Publ I. H. E. S. No. 58(1983), 197-212.

[405] RUAN W D. Canonical coordinates and Bergmann metrics[J]. Comm. Anal. Geom. 6(1998), no. 3, 589-631.

[406] RAY D R, SINGER I M. Analytic torsion for complex manifolds[J]. Ann. of Math. (2) 98(1973), 154-177.

[407] REID M. The moduli space of 3-folds with K = 0 may nevertheless be irreducible [J]. Math. Ann. 278(1987), no. 1-4, 329-334.

[408] ROSS J, THOMAS R P. An obstruction to the existence of constant scalar curvature Kähler metricsmath. arXiv: math. DG/0412518.

[409] ROSS J, THOMAS R P. A study of the Hilbert-Mumford criterion for the stability of projective varieties. arXiv: math. AG/0412519.

[410] ROYDEN H L, Report on the Teichmüler metric[J]. Proc. Nat. Acad. Sci. U.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

S. A.65(1970)497-499.

[411] RUAN Y B. Topological sigma model and Donaldson-type invariants in Gromov theory[J]. Duke Math.J.83(1996),no.2,461-500.

[412] RUAN Y B, TIAN G. A mathematical theory of quantum cohomology[J]. J. Differential Geom.42(1995),no.2,259-367.

[413] SACKS J, Uhlenbeck K. The existence of minimal immersions of 2-spheres[J]. Ann. of Math.(2)113(1981),no.1,1-24.

[414] SAPER L. On the cohomology of locally symmetric spaces and of their compactifications[M]. Somerville:Int. Press, 2003.

[415] SAPER L, STERN M. L_2-cohomology of arithmetic varieties[J]. Ann. of Math.,132(1990),no.1,1-69.

[416] SARNAK P. Additive number theory and Maass forms[M]. Berlin:Springer, 1984.

[417] SALOFF-COSTE L. Analysis on Riemannion co-compact covers. Survery in Differential Geometry, IX, 351-385[J]. International Press, Somerville, MA, 02143.

[418] SARNAK P. Spectra of hyperbolic surfaces[J]. Bull. Amer. Math. Soc.(N. S.)40(2003),no.4,441-478.

[419] SCHOEN R. Conformal deformations of a Riemannian metric to constant scalar curvature[J]. J. Differential Geom.20(1984),no.2,479-495.

[420] SCHOEN R, SIMON L, YAU S T. Curvature estimates for minimal hypersurfaces[J]. Acta Math. 134(1975),no.3-4,275-288.

[421] SCHOEN R, UHLENBECK K. A regularity theory for harmonic maps[J]. J. Differential Geom.17(1982),no.2,307-335.

[422] SCHOEN R, UHLENBECK K. Boundary regularity and the Dirichlet problem for harmonic maps[J]. J. Differential Geom.18(1983),no.2,253-268.

[423] SCHOEN R, WOLFSON J. Minimizing area among Lagrangian surfaces: the mapping problem[J]. J. Differential Geom.58(2001),no.1,1-86.

[424] SCHOEN R, YAU S T. On univalent harmonic maps between surfaces[J]. Invent. Math.44(1978),no.3,265-278.

[425] SCHOEN R, YAU S T. Existence of incompressible minimal surfaces and the topology of three-dimensional manifolds with nonnegative scalar curvature[J]. Ann. of Math.(2)110(1979),no. 1,127-142.

[426] SCHOEN R, YAU S T. On the proof of te positive mass conjecture in general relativity[J]. Comm. Math. Phys. 65(1979),no.1,45-76.





[427] SCHOEN R, YAU S T. Positivity of the total mass of a general space-time. Phys [J]. Rev. Lett. 43(1979), no. 20, 1457-1459.

[428] SCHOEN R, YAU S T. On the structure of manifolds with positive scalar curvature[J]. Manuscripta Math. 28(1979), no. 1-3, 159-183.

[429] SCHOEN R, YAU S T. Compact group actions and the topology of manifolds with nonpositive curvature[J]. Topology 18(1979), no. 4, 361-380.

[430] SCHOEN R, YAU S T. Proof of the positive mass theorem. II[J]. Comm. Math. Phys. 79(1981), no. 2, 231-260.

[431] SCHOEN R, YAU S T. Complete three-dimensional manifolds with positive Ricci curvature and scalar curvature[M]. Princeton: Princeton Univ. Press, 1982.

[432] SCHOEN R, YAU S T. The existence of a black hole due to condensation of matter[M]. Comm. Math. Phys. 90(1983), no. 4, 575-579.

[433] SCHOEN R, YAU S T. Conformally flat manifolds, Kleinian groups and scalar curvature[J]. Invent. Math. 92(1988), no. 1, 47-71.

[434] SCHOEN R, YAU S T. Lectures on differential geometry[M]. Cambridge: International Press, 1994.

[435] SEIBERG N, Witten W. Electric-magnetic duality, monopole condensation, and confinement in $N = 2$ supersymmetric Yang-Mills theory [J]. Nuclear Phys. B426(1994), no. 1, 19-52.

[436] SEIBERG N, WITTEN E. Monopoles, duality and chiral symmetry breaking in N = 2 supersymmetric QCD[J]. Nuclear Phys. B 431(1994), no. 3, 484-550.

[437] SEIDEL P, THOMAS R. Braid group actions on derived categories of coherent sheaves[J]. Duke Math. J. 108(2001), no. 1, 37-108.

[438] SHATAH J, STRUWE M. Regularity results for nonlinear wave equations[J]. Ann. of Math. (2)138(1993), no. 3, 503-518.

[439] SHI W X. Complete noncompact Kähler manifolds with positive holomorphic bisectional curvature[J]. Bill. Amer. Math. Soc. (N.S.)23(1990), no. 2, 437-440.

[440] SHI W X. RICCI flow and the uniformization on complete noncompact Kähler manifolds[J]. J. Differential Geom. 45(1997), no. 1, 94-220.

[441] SHI W X. A uniformization theorem for complete Kähler manifolds with positive holomorphic bisectional curvature [J]. J. Geom. Anal. 8 (1998), no. 1, 117-142.

[442] SHI Y G, TAM L F. Positive mass theorem and the boundary behaviors of compact manifolds with nonnegative scalar curvature [J]. J. Differential Geom. 62

(2002), no. 1, 79-125.

[443] SHIFFMAN B. Complete characterization of holomorphic chains of codimension one[J]. Math. Ann. 274(1986), no. 2, 233-256.

[444] SIMON L. Theorems on the regularity and singularity of minimal surfaces and harmonic mpas[J]. Geometry and global analysis(Sendai, 1993), 111-145.

[445] SIMON L. Existence of Willmore surfaces[M]. Canberra: Austral. Nat. Univ., 1986.

[446] SIMON L. Theorems on regularity and singularity of energy minimizing maps [M]. Basel: Birkhäuser Verlag, 1996.

[447] SIMPSON C T. Constructing variations of Hodge structure using Yang-Mills theory and applications to uniformization[J]. J. Amer. Math. Soc. 1(1988), no. 4, 867-918.

[448] SIMPSON C T. Moduli of representations of the fundamental grup of a smooth projective variet. I[J]. Inst. Hautes Études Sci. Publ. Math. No. 79(1994), 47-129.

[449] SIMPSON C T. Moduli of representations of the fundamental grup of a smooth projective Variety. II[J]. Inst. Hautes Études Sci. Publ. Math. No. 80 (1994), 5-79(1995).

[450] SINGER I M, WONG B, YAU S T, et al. An estimate of the gap of the first two eigenvalues in the Schrödinger operator[J]. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 12(1985), no. 2, 319-333.

[451] SIU Y T. Complex-analyticity of harmonic maps and strong rigidity of compact Kähler manifolds[J]. Ann. of Math. 112(1980), 73-111.

[452] SIU Y T. A simple proof of the surjectivity of the period map of K3 surfaces[J]. Manuscripta Math. 35(1981), no. 3, 311-321.

[453] SIU Y T. Every K3 surface is Kähler[J]. Invent. Math. 73(1983), no. 1, 139-150.

[454] SIU Y T. Effective very ampleness[J]. Invent. Math. 124(1996), no. 1-3, 563-571.

[455] SIU Y T. Invariance of plurigenera[J]. Invent. Math. 134(1998), no. 3, 661-673.

[456] SIU Y T. Very ampleness part of Fujita's conjecture and multiplier ideal sheaves of Kohn and Nadel[M]. Berlin: Inst. Publ., 2001.

[457] SIU Y T, YAU S T. Complete Kähler manifolds with nonpositive curvature of faster than quadratic decay[J]. Ann. of Math. (2) 105(1977), no. 2, 225-264.





[458] SIU Y T, YAU S T. Compact Kähler manifolds of positive bisectional curvature [J]. Invent. Math.59(1980), no.2, 189-204.

[459] SIU Y T, YAU S T. Compactification of negatively curved complete Kähler manifods of finite volume[M]. Princeton: Princeton Univ. Press, 1982.

[460] SMITH I, THOMAS R P, YAU S T. Symplectic conifold transitions[J]. J. Differential Geom. 62(2002), no.2, 209-242.

[461] SMOCZYK K. Longtime existence of the Lagrangian mean curvature folw[J]. Calc. Var. Partial Differential Equations 20(2004), no.1, 25-46.

[462] SMOCZYK K, WANG M T. Mean curvature flows of Lagrangians submanifolds with convex potentials[J]. J Differential Geom.62(2002), no.2, 243-257.

[463] SMOLLER J A, WASSERMAN A G, YAU S T. Existence of black hole solutions for the Einstein-Yang/Mills equations [J]. Comm. Math. Phys. 154 (1993), no.2, 377-401.

[464] SMOLLER J A, WASSERMAN A G, YAU S T, et al. Smooth static solutions of the Einstein/Yang-Mills equations[J]. Comm. Math. Phys. 143(1991), no.1, 115-147.

[465] Stern M. L^2-index theorems on locally symmetric spaces[J]. Invent. Math.96 (1989), no.2, 231-282.

[466] STOLZ S. Simply connected manifolds of positive scalar curvature[J]. Ann. of Math.(2)136(1992), no.3, 511-540.

[467] STROMINGER A. Superstrings with torsion[J]. Nuclear Phys. B 274(1986), no.2, 253-284.

[468] STROMINGER A, YAU S T, ZASLOW E. Mirror symmetry is T-duality[J]. Nuclear Phys. B 479(1996), no.1-2, 243-259.

[469] Stroock D W. An introduction to the analysis of paths on a Riemannian manifold [M]. Providence: American Mathematical Society, 2000.

[470] SULLIVAN D. A generalization of Milnor's inequality concerning affine foliations and affine manifolds[J]. Comment. Math. Helv.51(1976), no.2, 183-189.

[471] SULLIVAN D. The Dirichlet problem at infinity for a negatively curved manifold [J]. J. Differential Geom.18(1983), no.4, 723-732.

[472] SZEGÖ G G. Inequalities for certain eigenvalues of a membrane of given area [J]. J. Rational Mech. Anal.3(1954), 343-356.

[473] TALENTI G. Best constant in Sobolev inequality[J]. Ann. Mat. Pura Appl., 110(1976), 353-372.

[474] TAO T. Low-regularity global solutions to nonlinear dispersive equations[M].

Canberra: Austral. Nat. Univ., 2002.

[475] TATARU D. On global existence and scattering for the wave maps equation[J]. Amer. J. Math. 123(2001), no. 1, 37-77.

[476] TAUBES C H. Self-dual Yang-Mills connections on non-self-dual 4-manifolds [J]. J. Differential Geom. 17(1982), no. 1, 139-170.

[477] TAUBES C H. Self-dual connections on 4-manifolds with indefinite intersection matrix[J]. J. Differential Geom. 19(1984), no. 2, 517-560.

[478] TAUBES C H. S^1 actions and elliptic genera[J]. Comm. Math. Phys. 122 (1989), no. 3, 455-526.

[479] TAUBES C H. Casson's invariant and gauge theory[J]. J. Differential Geom. 31(1990), no. 2, 547-599.

[480] TAUBES C H. The existence of anti-self-dual conformal structures[J]. J Differential Geom. 36(1992), no. 1, 163-253.

[481] TAUBES C H. The Seibery-Witten invariants and symplectic forms[J]. Math. Res. Lett. 1(1994), no. 6, 809-822.

[482] TAUBES C H. SW⇒Gr: from the Seiberg-Witten equations to pseudoholomorphic curves[J]. J. Amer. Math. Soc. 9(1996), no. 3, 845-918.

[483] TAUBES C H. Counting pseudo-holomorphic submanifolds in dimension 4[J]. J Differential Geom 44(1996), no. 4, 818-893.

[484] TAUBES C H. Gr⇒SW: from pseudo-holomorphic curves to Seiberg-Witten solutions[J]. J. Differential Geom. 51(1999), no. 203-334.

[485] TAUBES C H. GR = SW: counting curves and connections[J]. J. Differential Geom. 52(1999), no. 3, 453-609.

[486] TAUBES C H. Seibery Witten and Gromov invariants for symplectic 4-manifolds [M]. Somerville: International Press, 2000.

[487] TERNG C L. A higher dimension generalization of the sine-Gordon equation and its soliton theory[J]. Ann. of Math. (2) 111(1980), no. 3, 491-510.

[488] TERNG C L. Recent progress in submanifold geometry[M]. Providence: Amer. Math. Soc., 1993.

[489] TERNG C L. Soliton equations and differential geometry[J]. J. Differential Geom. 45(1997), no. 2, 407-445.

[490] TERNG C L, UHLENBECK K. Poisson actions and scattering theory for integrable systems[M]. Boston: Int. Press, 1998.

[491] TERNG C L, UHLENBECK K. Bäcklund transformations and loop group actions [J]. Comm. Pure Appl. Math. 53(2000), no. 1, 1-75.



[492] THOMAS R P, YAU S T. Special Lagrangians, stablebundles and mean curvature flow[J]. (English. English summary) Comm. Anal. Geom. 10(2002), no. 5, 1075-1113.

[493] THORBERGSSON G. A survey on isoparametric hypersurfaces and their generalizations[M]. Amsterdam: North-Holland, 2000.

[494] THURSTON W P. Three-dimensional manifolds, Kleinian groups and hyperbolic geometry[J]. Bull. Amer. Math. Soc. (N.S.) 6(1982), no. 3, 357-381.

[495] THURSTON W. Three-dimensional Geometry and Topology, Vol. 1[M]. Princeton: Princeton Univ. Press, 1997.

[496] TIAN G. Smoothness of the universal deformation space of compact Calabi-Yau manifolds and its Petersson-Weil metric[M]. Singapore: World Sci. Publishing, 1987.

[497] TIAN G. On a set of polarized Kähler metrics on algebraic manifolds[J]. J. Differential Geom. 32(1990), no. 1, 90-130.

[498] TIAN G. Calabi's conjecture for complex surfaces with positive first Chern class [J]. Invent. Math. 101(1990), no. 1, 101-172.

[499] TIAN G. Kähler-Einstein metrics with positive scalar curvature[J]. Invent. Math. 130(1997), no. 1, 1-37.

[500] TIAN G. Gauge theory and calibrated geometry. I[J]. Ann. of Math., 151 (2000), no. 1, 193-268.

[501] TIAN G, YAU S T. Three-dimensional algebraic manifolds with $C_1 = 0$ and $\chi = -6$[M]. Singapor: World Sci. Publishing, 1987.

[502] TIAN G, YAU S T. Complete Kähler manifolds with zero Ricci curvature. I[J]. J. Amer. Math. Soc 3(1990), no. 3, 579-609.

[503] TIAN G, YAU S T. Complete Kähler manifolds with zero Ricci curvature. II [J]. Invent. Math. 106(1990), no. 1, 27-60.

[504] TODOROV A N. Applications of the Kähler-Einstein-Calabi-Yau metric to moduli of K3 surfaces[J]. Invent. Math. 61(1980), no. 3, 251-265.

[505] TODOROV A N. The Weil-Petersson geometry of the moduli space of SU($n \geqslant 3$) (Calabi-Yau) manifolds. I[J]. Comm. Math. Phys. 126(1989), no. 2, 325-346.

[506] TREIBERGS A E. Entire spacelike hypersurfaces of constant mean curvature in Minkowski space[J]. Invent. Math. 66(1982), no. 1, 39-56.

[507] TRUDINGER N S. Remarks concerning the conformal deformation of Riemannian structures on compact manifolds[J]. Ann. Scuola Norm. Sup. Pisa (3) 22 (1968), 265-274.

[508] TRUDINGER N S, WANG X J. The Bernstein problem for affine maximal hypersurfaces[J]. Invent. Math. 140(2000), no. 2, 399-422.

[509] UHLENBECK K K. Removable singularities in yang-Mills fields[J]. Comm. Math. Phys. 83(1982), no. 1, 11-29.

[510] UHLENBECK K K. Connections with L^p bounds on curvature[J]. Comm. Math. Phys. 83(1982), no. 1, 31-42.

[511] UHLENBECK K. Harmonic maps into Lie groups: classical solutions of the chiral model[J]. J. Differential Geom. 30(1989), no. 1, 1-50.

[512] UHLENBECK K, YAU S T. On the existence of Hermitian-Yang-Mills connections in stable vector bundles[J]. Comm. Pure Appl. Math. 39(1986), no. S, suppl., S257-S293.

[513] VAFA C. Topological mirrors and quantum rings[M]. Hong Kong: Internat. Press, 1992.

[514] VERDIERE C. Theoremes de Courant et de Cheng Combinatoires. Combinatorial Courant and Cheng theorems Seminaire de Theorie Spectrale et Geometrie, No. 13, Annee 1994-1995, 9-13.

[515] WAN T. Constant mean curvature surface, harmonic maps, and universal Teichmüller space[J]. J. Differential Geom. 35(1992), no. 3, 643-657.

[516] WANG C P. Some examples of complete hyperbolic affine 2-spheres in R^3[M]. Berlin: Springer, 1991.

[517] WANG M T. A fixed point theorem of discrete group actions on Riemannian manifolds[J]. J. Differential Geom. 50(1998), no. 2, 249-267.

[518] WANG M T. Generalized harmonic maps and representations of discrete groups [J]. Comm. Anal. Geom. 8(2000), no. 3, 545-563.

[519] WANG M T. Mean curvature flow of surfaces in Einstein four-manifolds[J]. J. Differential Geom. 57(2001), no. 2, 301-338.

[520] WANG M T. Deforming area preserving diffeomorphism of surfaces by mean curvature flow[J]. Math. Res. Lett. 8(2001), no. 5-6, 651-661.

[521] WANG M Y, ZILLER W. Existence and nonexistence of homogeneous Einstein metrics[J]. Invent. Math. 84(1986), no. 1, 177-194.

[522] WANG X J, ZHU X H. Kähler-Ricci solitons on toric manifolds with positive first Chern class[J]. Adv. Math. 188(2004), no. 1, 87-103.

[523] WANG X W. Balance point and stability of vector bundles over a projective manifols[J]. Math. Res. Lett. 9(2002), no. 2-3, 393-411.

[524] WEINBERGER. S. Aspects of the Novikov conjecture, in Geometric and Topo-





logical Invariants of Elliptic Operators[M]. Providence: AM. S., 1990.

[525] WICKRAMASEKERA N. A rigidity theorem for stable minimal hypercones[J]. J. Differential Geom. 68(2004), no. 3, 433-514.

[526] WICKRAMASEKERA N. On the singularities and a Bernstein property of immersed stable minimal hypersurfaces. Calc. Var. Partial Differential Equations 22(2005), no. 1, 1-20.

[527] Wtten E. A new proof of the positive energy theorem[J]. Comm. Math. Phys. 80(1981), no. 3, 381-402.

[528] WITTEN E. Supersymmetry and Morse theory [J] J. Differential Geom. 17 (1982), no. 4, 661-692(1983).

[529] WITTEN E. Holomorphic Morse inequalities[M]. Leipzig: Teubner-Texte Math., 1984.

[530] WITTEN E. Elliptic genera and quantum field theory[J]. Comm. Math. Phys. 109(1987), no. 4, 525-536.

[531] WITTEN E. 2 + 1-dimensional gravity as an exactly soluble system[J]. Nuclear Phys. B 311(1988/89), no. 1, 46-78.

[532] WITTEN E. On the structure of the topological phase of two-dimensional gravity [J]. Nuclear Phys. B 340(1990), no. 2-3, 281-332.

[533] WITTEN E. Monopoles and four-manifolds[J]. Math. Res. Lett. 1(1994), no. 6, 769-796.

[534] WITTEN E, YAU S T. Connectedness of the boundary in the AdS/CEF correspondence[J]. Adv. Theor. Math. Phys. 3(1999), no. 6, 1635-1655(2000).

[535] WOLF W. The Teichmuller theory of harmonic maps[J]. J. Differential Geom. 29(1989), no. 2, 449-479.

[536] WOLPERT S A. Geometry of the Weil-Petersson completion of Teichmüller space [M]. Somerville: Int. Press, 2003.

[537] YANG P C, YAU S T. Eigenvalues of the Laplacian of compact Riemann surfaces and minimal submanifolds[J]. Ann. Scuola Norm. Sup. Pisa Cl. Sci., 7 (1980), no. 1, 55-63.

[538] YAU S T. On the fundamental group of compact manifolds of non-positive curvature. Ann. of Math. (2)93(1971), 579-585.

[539] YAU S T. Harmonic functions on complete Riemannian manifolds[J]. Comm. Pure Appl. Math., 28(1975), 201-228.

[540] YAU S T. Isoperimetric constants and the first eigenvalue of a compact Riemannian manifold[J]. Ann. Sci. École Norm. Sup., 8(1975), no. 4, 487-507.

[541] YAU S T. Some function-theoretic properties of complete Riemannian manifold and their applications to geometry[J]. Indiana Univ. Math.J.25(1976),no.7, 659-670.

[542] YAU S T. Calabi's conjecture and some new results in algebraic geometry[J]. Proc. Nat. Acad. Sci. U.S.A.74(1977),no.5,1798-1799.

[543] YAU S T. On the Ricci curvature of a compact Kähler manifold and the complex Monge-Ampère equation. I[J]. Comm. Pure Appl. Math.31(1978),no.3,339-411.

[544] YAU S T. The role of partial differential equations in differential geometry[M]. Helsinki:Acad. Sci. Fennica, 1980.

[545] YAU S T. A general Schwarz lemma for Kähler manifolds[J]. Amer. J. Math. 100(1978),no.1,197-203.

[546] YAU S T. Problem section. Seminar on Differential Geometry[M]. Princeton: Princeton Univ. Press, 1982.

[547] YAU S T. Compact three-dimensional Kähler manifolds with zero Ricci curvature [M]. Singapore:World Sci. Publishing, 1985.

[548] YAU S T. Nonlinear analysis in geometry [M]. Geneva: L'Enseignement Mathématique, 1986.

[549] YAU S T. A review of complex differential geometry[J]. Proc. Symp. Pur Math.52(1991),619-625.

[550] YAU S T. Open problems in Geometry[J]. Proc. Symp. Pure Math.54(1933) 1-28.

[551] YAU S T. An application of eigenvalue estimate to algebraic curves defined by congruence subgroups[J]. Math. Res. Lett.3(1996),no.2,167-172.

[552] YAU S T. A note on the distribution of critical points of eigenfunctions[M]. River Edge:World Sci. Publishing, 1997.

[553] YAU S T. An estimate of the gap of the first two eigenvalues in the Schrödinger operator[M]. Somerville:Int. Press, 2003.

[554] YAU S T, MEEKS W H. The equivariant loop theorem for three-dimensional manifolds and a review of the existence theorems for minimal surfaces[M]. Orlando:Acadmic Press, 1984.

[555] YE R G. Gromov's compactness theorem for pseudo holomorphic curves[J]. Trans. Amer. Math. Soc.342(1994),no.2,671-694.

[556] YE R G. Foliation by constant mean curvature spheres on asymptotically flat manifolds[M]. Cambridge:Internat. Press, 1996.



[557] YU G L. Novikov conjecture for groups with finite asymptotic dimension[J]. Ann. of Math. (2) 147(1998), no. 2, 325-355.

[558] ZELDITCH S. Szegö kernels and a theorem of Tian[J]. Int. Math. Res. Not. 6 (1998), 317-331.

[559] ZELDITCH S. Spectral determination of analytic bi-axisymmetric plane domains [J]. Geom. Funct. Anal. 10(2000), no. 3, 628-677.

[560] ZHANG S W. Heights and reductions of semi-stable varieties[J]. Compositio Math. 104(1996), no. 1, 77-105.

[561] ZHANG X. A definition of total energy-momenta and the positive mass theorem on asymptotically hyperbolic 3-manifolds. I [J]. Comm. Math. Phys. 249 (2004), no. 3, 529-548.

[562] ZHONG J Q, YANG H C. On the estimate of the first eigenvalue of a compact Riemannian manifold[J]. Sci. Sinica Ser. A 27(1984), no. 12, 1265-1273.

[563] ZINGER A. Enumeration of genus-two curves with a fixed complex structure in P^2 and P^3, arXiv: math. SG/0201254.

[564] ZINGER A. Enumerative vs. Invariants and obstruction bundles, arXiv: math. SG/0201255.

附录Ⅸ The Excerpts from the Geometric Topology of 3-Manifolds

CHAPTER Ⅱ PL PLANAR MAPS①

There is merit in moving rapidly through the first two chapters. The theorems are for background and are useful tools. Our familiarity with the plane causes us to introduce many theorems in that space. In most cases the theorems extend to higher dimensions by the same techniques, and we will not repeat the proofs.

Ⅱ1. Linear maps. A continuous function or transformation is called a *map*. A map of simplex Δ into Euclidean space (or another simplex) is called *linear* if f preserves linearity and ratios. We say that *f preserves linearity* if straight-line segments in Δ are sent by f into straight-line segments (or points). It *preserves ratios* if for each three collinear points x_1, x_2, x_3 of Δ, $f(x_2)$ divides $f(x_1), f(x_3)$ in the same ratio that x_2 divides x_1, x_3—that is, $d(x_1, x_2)\cdot d(f(x_2), f(x_3)) = d(x_2, x_3)\cdot d(f(x_1), f(x_2))$ where d denotes Euclidean distance. Projection maps preserve linearity, but some do not preserve rations and are hence not linear. Figure Ⅱ.1 shows the projection of triangle abc into a plane π from a point p. It is not true that a point halfway between a and b projects into a point halfway between a' and b'.

The above definition of a linear map is geometric. There is also the physical approach where a map f of a simplex into Euclidean m-space R^n is defined to be linear if the center of gravity of weights hung onto the vertices of the simplex is sent by f to the center of gravity of the same weights hung on the corresponding images of the vertices. We shall not use this physical approach but mention it for its visual appeal.

There is also the algebraic approach where we recall from algebra that $a_0 + a_1x_1 + \cdots + a_nx_n$ is a linear function on n variables (the a's are constant and the x's variables). It is convenient to denote a point x of R^n by its coordinates $(x_1, x_2, \cdots, x_n)$ so we write $x = (x_1, x_2, \cdots, x_n)$. A map f of a subset X of R^n into R^n is defined to be linear if there are m linear functions $f_1(x) = a_{i0} + a_{i1}x_1 + \cdots + a_{il}x_n$ $(i = 1, 2, \cdots, m)$ such

① 转载自 R. H. Bèng.《三维流形的几何拓扑》.



that for each $x \in X, f(x)$ has coordinates $f_1(x), f_2(x), \cdots, f_m(x)$.

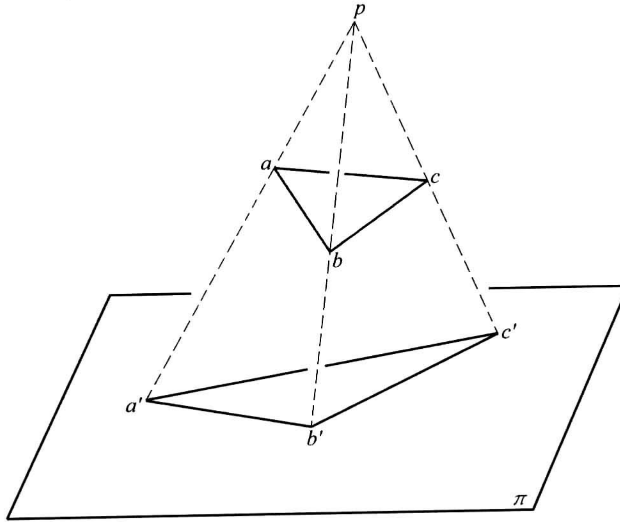


Figure Ⅱ.1

If p_1, p_2 are planar points with coordinates $(x_1, y_1), (x_2, y_2)$ and a_1, a_2 are constants, we use $a_1p_1 + a_2p_2$ to denote the point with coordinates $(a_1x_1 + a_2x_2, a_1y_1 + a_2y_2)$. Each point $p \in v_1v_2v_3$ can be written as $a_1v_1 + a_2v_2 + a_3v_3$ where a_1, a_2, a_3 are nonnegative numbers whose sum is 1. Exercise Ⅱ.1.A shows one way to find a_1, a_2, a_3 if v_1, v_2, v_3, p are given.

EXERCISE Ⅱ.1.A. Suppose f is a map of triangular disk $v_1v_2v_3$ into triangular disk $v'_1v'_2v'_3$ described as follows: $f(v_i) = v'_i (i = 1,2,3)$; if x is between v_2 and v_3, $f(x)$ is the point between v'_2 and v'_3 that divides v'_2, v'_3 is the same ratio that x divides v_2, v_3; if x is between v_2 and v_3 and y is between v_1 and x, then f(y) is the point between v'_1 and $f(x)$ that divides v'_1, $f(x)$ in the same ratio that y divides v_1, x. Show that if the ratios are $t:1-t$ and $s:1-s$ respectively as suggested by Figure Ⅱ.1.A and the points are planar vectors, then

$$x = tv_3 + (1-t)v_2$$

$$y = (1-s)v_1 + stv_3 + s(1-t)v_2, \text{and}$$

$$f(y) = (1-s)v'_1 + stv'_3 + s(1-t)v_2$$

While one could use equations to show that various definitions of linearity are equivalent, it is to be admitted that geometric topologists are more concerned with concepts that with equations. A frequently used basic concept is that under a linear map the image of a simplex is the convex union of a finite number of simplexes.

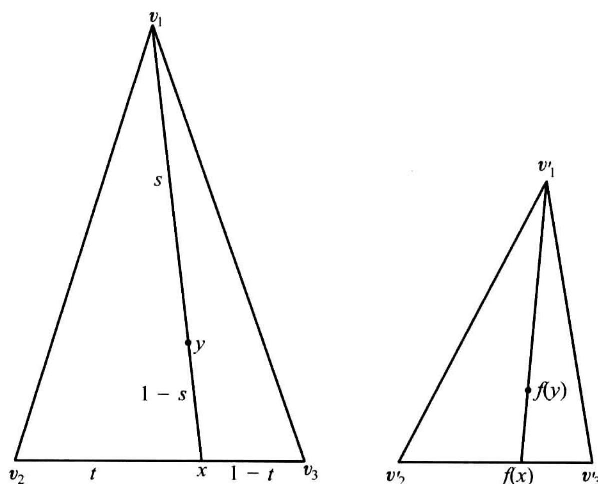


Figure Ⅱ.1.A

Ⅱ 2. PL maps. If (X, T) is a complex, we call a map f of X into R^2(or R^n) a *linear* map of(X, T) if f is linear on each simplex of T. A map of a complex (X, T) into complex$(Y, S)$is *linear* if f takes each simplex of T linearly *into* a simplex of S. (Some authors insist that f takes each simplex of T linearly *onto* a simplex of S but we do not.) A map f of a complex(X, T)into a complex(Y, S) is *piecewise linear* (PL)if there is a subdivision t_1 of T such that f takes each simplex of T_1 linearly into a simplex of S.

THEOREM Ⅱ.2.A. *Suppose ϕ is a linear map of a 2-simplex Δ into R^2 and T_2 is a triangulation of $\phi(\Delta)$ such that the image of each face of Δ is the union of elements of T_2. Then there is a triangulation T of Δ such that ϕ takes each simplex of T linearly onto a simplex of T_2.*

PROOF. If $\phi(\Delta)$ is a point, any subdivision of Δ would serve for T, but to be specific, we pick the one where $\Delta \in T$.

If ϕ is a homeomor-phism, we describe T by regurring that for each $\sigma \in T_2$, $f^{-1}(\sigma) \in T$.

If $\phi(\Delta)$ is of dimension 1, the inverses of 1-simplexes of T_2 are triangular or quadrilateral disks as shown in Fgure Ⅱ.2.A and the four-sided subdisks are subdivided as suggested by the dotted lines.

THEOREM Ⅱ.2.B. *If(X, T) is a compact 2-complex and f is a linear map of (X, T) into R^2, then there is a triangulation T_2 of R^2 and a subdivision T_1 of T such that f takes each simplex of T_1 linearly onto a simplex of T_2.*

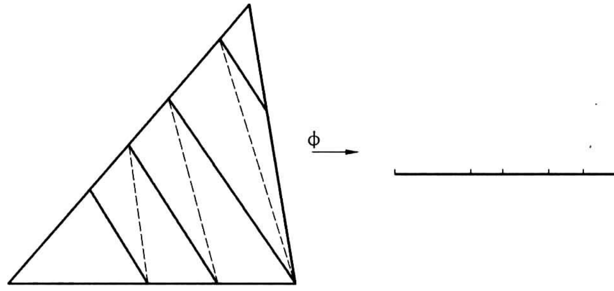


Figure Ⅱ.2.A

PROOF. Let T_2 be a triangulation of R^2 sch that the image under f of each simplex of T is the union of simplexes of T_2. Any refinement of T_2 would serve for the T_2 of the statement of the theorem but we use T_2 itself.

We get T_1 by subdividing each simplex Δ of T. We must not subdivide elements of T too much lest their images be proper subsets of elements of T_2. If $f(\Delta)$ is a vertex of T_2, $\Delta \in T_1$. If Δ^1 is a 1-simplex and f is a homeomorphism on Δ^1, f^{-1} on the triangulation T_2 of $f(\Delta^1)$ provides a subdivision of Δ^1. If Δ^2 is a 2-simplex of T, Theorem Ⅱ.2.A provides a subdivision of Δ^2.

EXAMPLE Ⅱ.2.C. Theorem Ⅱ.2.B is false if one leaves out the hypothesis that X is compact. Consider the 1-dimensional case where X is the ray $[0, \infty]$ and T is a triangulation of X whose vertices are integral points. Let f take the segment $[i-1, i]$ linearly onto the segment $[1/i, 1/(i+1)]$. There is no subdivision T_1 of T and triangulation T_2 of R^2 such that f takes each simplex of T_1 *onto* a simplex of T_2.

THEOREM Ⅱ.2.D. *If f is a linear map of a 2-complex (X_1, T_1) into a 2-complex (X_2, T_2) and S_2 is a subdivision of T_2, then there is a subdivision S_1 of T_1 so that f takes (X_1, S_1) linearly into (X_2, S_2).*

PROOF. We explain how to subdvide T_1 to get S_1. We do not subdivide a simplex Δ of T_1 if $f(\Delta)$ lies in an element of S_2.

If Δ^1 is a 1-simplex of T_1 and f is a homeomorphism on Δ^1, we get a description of S_1 on Δ^1 by subdividing Δ^1 so that the image under f of each of the pieces lies in an element of S_2. Similarly, if Δ^2 is a 2-simplex of T_1 and $f(\Delta^2)$ is of dimension 1, we have a situation like that depicted in Figure Ⅱ.2.A and can subdivide Δ^2 in a manner as suggested there. If f is a homeomorphism on 2-simplex Δ^2 of T, we get a triangulation T' of (Δ^2) such that if σ^1 is a 1-simplex of S_2 intersecting $f(\Delta^2)$. $\sigma^1 \cap f(\Delta^2) \in T'$. This is done without putting unnecessary vertices on Bd $f(\Delta^2)$. The elements of S_1 in Δ^2 are the inverses of the elements of T'.

THEOREM Ⅱ.2.E. *If $(X_1, T_1), (X_2, T_2), (X_3, T_3)$ are 2-complexes, f_1 is a PL map from (X_1, T_1) into (X_2, T_2), and f_2 is a PL map from (X_2, T_2) into (X_3, T_3) then $f_2 f_1$ is a PL map of (X_1, T_1) into (X_3, T_3).*

PROOF. Since f_2 is PL there is a refinement T'_2 of T_2 such that $f_2:(X_2, T'_2)\to(X_3, T_3)$ is linear. Also there is a refinement T'_1 of T_1, such that $f_1:(X_1, T'_1)\to(X_2, T_2)$ is linear. Since T'_2 is a refinement of T_2, we find from Theorem Ⅱ.2.D that there is a refinement T''_1 of T_1 such that $f_1:(X_1, T''_1)\to(X_2, T'_2)$ is linear. Hence f_2f_1 takes each simplex of T''_1 int a simplex of T_3.

Since the composition of linear maps is linear, f_2f_1 is.

Ⅱ.3. Pushes. We shall describe a special useful kind of homeomorphism of R^2 onto itself called a *push* in R^2. Suppose v_0 is an arbitrary vertex of an arbitrary triangulation T of R^2. The *star of* v_0 (denoted by Star)(v_0, T) or merely by Star v_0 if no confusion results is the union of the simplexes of T containing v_0. The *link of* v_0 (denoted by Link v_0) is the union of all 1-simplexes v_iv_j of T such that $v_0v_iv_j$ is a 2-simplex of T.

A set X can be *starred* from a point $x\in X$ if each closed ray from x intersects X in a connected set. Note that the PL disk Star v_0 can be starred from v_0, and each ray from v_0 intersects the polygon Link v_0 = Bd Star v_0 in just one point.

Suppose $p, p'\in R^2$ and T, T' are rectilinear triangulations of R^2 such that Star p in T is the same set as Star p' in T'. The homeomorphism of R^2 onto itself that is fixed outside Star p, takes p to p′, and takes the simplexes of T in Star p lineraly onto simplexes of T′ is called a *push*. Usually we regard the use to p to p′ as a function of p, p' but not necessarily of T, T'. There are many pushes of p *of* p' depending on T's used. We think of a push as the end of a 1-parameter family of linear homeomorphisms of (R^2, T), the tth of which sends p to $(1-t)p+tp'$.

A map h of a set X onto itself has *compact support* if X contains a compact set C such that h is the identity except possibly on C.

Note that a push in R^2 is a PL homeomorphism with compact support.

THEOREM Ⅱ.3.A. *The composition of a finite number of pushes in R^2 is a PL homeomorphism with compact support.*

Two complexes $(X_1, T_1), (X_2, T_2)$ in R^2 are in *general position* there if the 0-skeletons of neither intersects either the 0-skelton or 1-skelton of the other. If a 1-simplex of T_1 intersects a 1-simplex of T_2, they cross. Pushes provide a convenient way to shift one complex so that it becomes in general position with respect to another. Additional consideration of general position is found in § Ⅳ.7. For example. (X_1, T_1), (X_2, T_2) are in general position in R^3 if no vertex of $T_1 (i=1,2)$ lies on a vertex, 1-





simplex, or 2-simplex of $T_1(j \neq i)$, and no 1-simplex of T_1 intersects any 1-simplex of T_2.

A finite complex (M, T) is called a compact 2-*manifold-with-boundarv* if the star of each vertex of T is a disk.

THEOREM Ⅱ.3.B. *Suppose in* R^3 *that* $(M_1, T_1), (M_2, T_2)$ *are compact manifolds-with-boundaries which are in general position. The each componen* $M_1 \cap M_2$ *is either an arc or a simple closed curve.*

THEOREM Ⅱ.3.C *Suppose* f_1, f_2 *are linear maps of compact* 2-*manifolds-with boundaries* (M_1, T_1), (M_2, T_2) *into* R^3 *such that each* f_i *is a homeomorphism a* $f_1(T_1^i) \cap f_2(T_2^j) \neq \emptyset$ *if* $i + j < 3$. *Then each component of* $f_1^{-1}(f_1(M_1) \cap f_2(M_2))$ *is either an arc or a simple closed curve.*

If a set X and its image $f(X)$ lies in the same metric space, we say that $d(f, \mathrm{Id}) \leqslant \varepsilon$ if for each $x \in X$, the distance $d(f(x), x)$ between a point x and its image $f(x)$ is no more than ε. We say that $d(f, \mathrm{Id}) < \varepsilon$ if there is a number δ such that $d(f, \mathrm{Id}) < \delta \leqslant \varepsilon$.

THEOREM Ⅱ.3.D. *Suppose* $(X_1, T_1), (x_2, T_2)$ *are two rectilinear complexes is* R^2 *and* ε *is a positive number. Then if* X_1 *is compact, there is a PL homeomorphism* h $R^2 \rightarrow R^2$ *such that* h *is linear on* (X_1, T_1), $d(h, \mathrm{Id}) < \varepsilon$, h *has compact support, and* $(X_2, T_2), (h(X_1), h(T_1))$ *are in general position where* $h(T_1)$ *denotes the set images under* h *of the simplexes of* T_1.

PROOF. Let $v_1, v_2, \cdots, v_n$ be the vertices of T_1 and T be a triangulation of R^2 such that $T_1 \subset T$. See Theorem Ⅰ.1.B. We shall get a homeomorphism h by pushing v_1, then pushing $v_2 \ldots$, and finally pushing v_n. Each push will be less than ε.

To get the first push we get a point v'_1 within ε of v_1 such that Star(v_1, T) is starred with respect to v'_1 and v'_1 does not lie on any straight line determined by pairs of vertices in $T_1 \cup T_2$. The push h from v_1 to v'_1 changes the 1-simplexes of T_1 to 1-simplexes that are in general position with respect to (X_2, T_2). For simplicity we suppose h was the identity so as to be able to refer to T_1 instead of $h(T_1)$.

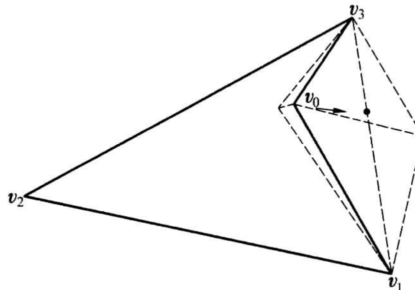


Figure Ⅱ.3.E

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

The proof is continued by pushing v_2, then $v_3, \ldots$, and finally v_n.

It is to be noted that we could have dropped the condition that X_1 is compact if we had not required that h have compact support.

The next two exercises along with the theorem which follows are included as preliminaries to Chapter Ⅲ.

EXERCISE Ⅱ.3.E. If $v_0v_1v_2v_3$ is a concave quadrilateral as shown in Figure Ⅱ.3.E, there is a push h that sends $v_0v_1v_2v_3$ to a triangle.

A suitable push sends v_0 to some point of Int $v_1v_3 = (v_1v_3)$. This exercise is included not because it is significant or interesting but rather because it is related to the important PL Schoenflies theorem discussed in the next chapter. The following exercise is another variation of the same idea.

EXERCISE Ⅱ.3.F. If $v_1v_2v_3v_4v_5$ is a regular polygon as shown in Figure Ⅱ.3.F, there is a push that sends $v_1v_2v_3v_4v_5$ to a quadrilateral.

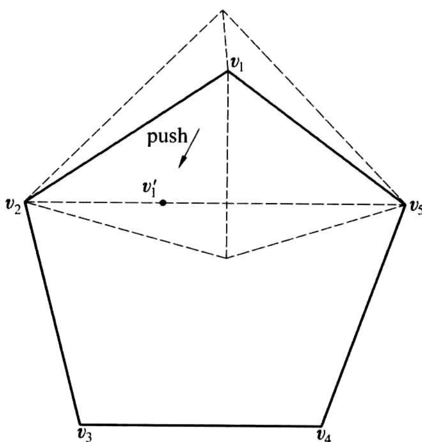
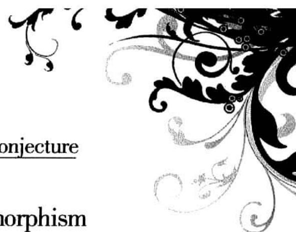


Figure Ⅱ.3.F

By pushing the vertices of a triangle, one gets the following result which is used in the next chapter.

THEOREM Ⅱ.3.G. *Suppose D is a triangular disk in R^2 and x, y, z are three noncollinear points of* Bd D, *then there are pushes h_1, h_2, h_3 on the vertices of D such that $h_3h_2h_1(D)$ is a triangular disk with vertices x, y, z.*

Ⅱ.4. **Isotopies.** Suppose $h: X \times [0,1] \to Y$ is continuous. We may denote $H(x \times t)$ by $H_t(x)$ and call $H_t(0 \leqslant t \leqslant 1)$ a 1-parameter family of maps of X into Y. A member of the family $\{H_t\}$ is a map H_t of X where $H_t(x) = H(x \times t)$. We shall not try to define a distance between elements of $\{H_t\}$. Note that X not be compact. We think of the family $\{H_t\}$ as being continuous sind $X \times [0,1] \to Y$ is continuous.



An *isotopy* of a space X into a space Y is a 1-parameter family of homeomorphism $H_t(0\leqslant t\leqslant 1)$ of X into Y. It is required that $\{H_t\}$ be simultane continuous in t and X rather than that the homeomorphisms be near each one. For example, a rotation of R^2 about the origin generates an isotopy even each pair of the homeomorphisms are infinitely far apar. (Some mathematical prefer to regard an isotopy as a map H of $X\times[0,1]$) into Y such that H restricted to each $X\times t$ is a homeomorphism. This concept is equivalent to the one use In many cases we will be considering the case where $Y=X$, each H_t is onto, one of H_0, H_1 is the identity. One can think of such an isotopy as a motion of X onto itself using t as a time variable. A push is such an isotopy.

The following theorem generalizes to all dimensions.

THEOREM Ⅱ.4.A. *If h is a homeomorphism of a disk D onto itself that is fixea* Bd D, *then there is an isotopy $H_t(0\leqslant t\leqslant 1)$ of D onto itself such that $H_0=H_1=$ identity, and each H_t is fixed on* Bd D.

PROOF. There are several proofs of this result, but we think of the "mgrinder" approach suggestd by the idea that when material is forced through a sieve or meat grinder, it comes out in smooth threads. Suppose that one starts with D as distorted by h and straightens it out by shrinking Bd D toward the center of D. See Figure Ⅱ.4.A, where D is regarded as a rectangle, and we show three views of it. If at time t, Bd D has been squeezed to the boundary of the smaller rectangle shown in the upper left part of Figure Ⅱ.4.A, then H_t is the identity in the part of D outside this rectangle. To describe H_t on the small rectangle, one might use a linear homeomorphism to expand the smaller rectangle and make it fit onto the larger, then use $H_0=h$ to send this larger rectangle on the upper right version of D, and finally use a linear homeomorphism to send the ectangle at the upper right of **Figure Ⅱ.4.A** onto the small rectangle in the lower center version of D.

A homeomorphism h on X is called an ε-homeomorphism if $d(x,h(x))\leqslant\varepsilon$ for each $x\in X$. The following theorem generalizes easily to all dimensions.

THEOREM Ⅱ.4.B. *Suppose $\varepsilon>0$ and h is an ε-homeomorphism of R^2 onto itself Then there is an isotopy $H_t(0\leqslant t\leqslant 1)$ of R^2 onto itself such that $H_1=h$ $H_0=$ identity, and each H_t is an ε-homeomorphism.*

Recall that for each point $p=(x,y)\in R^2$, tp is the point with coordinate(tx, ty). Let g_t be the homeomorphism of R^2 onto itself that sends each point p to tp. In proving Theorem Ⅱ.4.B, J.M. Kister[K3]made use of the function $g_1hg_{1/t}$. Note that if $0<t<1$, then $g_thg_{1/t}(p)=H_t(0<t<1)$ is obtained by pushing p away from the origin, then moving it by h, and finally pulling it toward the origin. One uses the hy-

pothesis that h is an ε-homeomorphism to show $H_0 = \lim_{t\to 0} g_t h g_{1/t}$ is the identity.

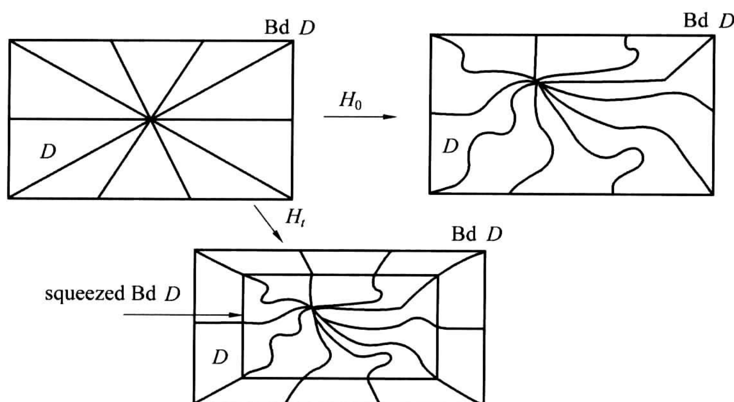


Figure Ⅱ.4.A

Ⅱ.5. Meshing triangulations. Suppose as shown in Figure Ⅱ.5 that U_1, U_2 are open sets with curvilinear triangulations T_1, T_2 which do not agree on $U_1 \cap U_2$. We seek a triangulation T of $U_1 \cup U_2$ that is compatible with T_1 on $U_1 - \overline{U_2}$ and with T_2 on $U_2 - \overline{U_1}$. We say that two triangulations S_1, S_2 of X_1, X_2 are *locally compatible* at a point $p \in X_1 \cap X_2$ if there are subdivisions S'_1, S'_2 of S_1, S_2 such that the identity map takes the union of the simplexes of S'_1 containing p linearly onto the union of the simplexes of S'_2 containing p. Also, S_1 and S_2 are *compatible* on $X_1 \cap X_2$ if there is a triangulation T of $X_1 \cap X_2$ which is subdivision of both S_1 and S_2.

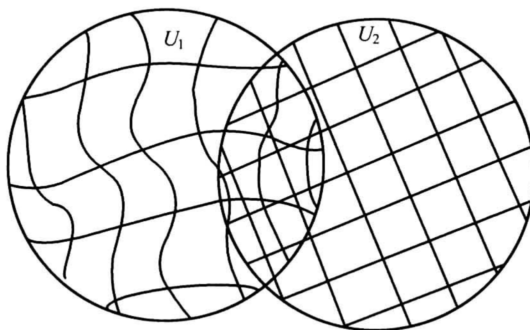


Figure Ⅱ.5

If T is a rectilinear triangulation of an open subset of R^2, then any 1-simplex of T is an edge of two and only two 2-simplexes of T. It follows from Chapter I that curvilinear triangulations of open subsets of R^2 also have this property, and we tacitly assume it in the next theorem, a variation of which can be used to show that the union of certain triangulated sets can be triangulated.





THEOREM. Ⅱ.5. *Suppose*

U_1, U_2 are two open subsets of R^2, $T_i(i=1,2)$ is a curvilinear triangulation of U_i, B is the union of 1-simplexes Δ^1 of T_1 such that one but not both of the 2-simplexes of T_1 with Δ^1 as an edge lies in U_2, h is a homeomorphism of R^2 onto itself that is fixed off $U_1 \cap U_2$ and takes each 1-simplex of T_1 in B onto a 1-simplex of T_2.

Then there is a curvilinear triangulation T of $U_1 \cup U_2$ such that T is compatible with T_1 on $U_1 - \overline{U_2}$ and with T_2 on $U_2 - \overline{U_1}$.

PROOF. Let M be union of 2-simplexes of T_1 that intersect $U_1 - U_2$. Note that $B = \text{Bd } M$ in U_1. If Δ is a simplex of T_1 in M, $h(\Delta)$ is an element of T. If Δ is a simplex of T_2 not in $h(M)$, $h(M) \cap \text{Int } \Delta = \emptyset$. Each simplex of T_2 not in $h(M)$ is an element of T.

It remains to assign $U_1 \cup U_2$ a metric so that each simplex of T is isometric to a triangular disk. We shall do this so that each is isometric to a unit equalateral triangular disk D_0^2. It is easy to get a global metric if we metrize each simplex of T so we concentrate on the elements of T.

If Δ^2 is a simplex of T_1 in M, there is a linear homeomorphism g of D_0^2 onto Δ^2 under T_1, and we pick a metric for $h(\Delta^2)$ of T by deciding that $hg: D_0^2 \to h(\Delta^2)$ is an isometry. As a first approximation to metrics on other 2-simplexes of T, we pick a metric for a 2-simplex $\Delta^2 \in T$ in $U_2 - h(\text{Int } M)$ by letting g be a linear homeomorphism of D_0^2 onto Δ^2 under T_2 and deciding that g is an isometry. The defect in this metric is that if Δ^1 is a 1-simplex of T_1 in B, we have assigned two metrices to $h(\Delta^1)$. The situation is remedied by considering a 2-simplex Δ^2 of T_2 that does not lie in $h(M)$ but contains an edge Δ^1 in $h(B)$ and modifying the metric on Δ^2. Let f be a homeomorphism Δ^1 onto itself that is fixed on Bd Δ^1 and is an isometry in taking the second metric (one associated with T_2) into the first. Let p be a point in Int Δ^2 so near Δ^1 that the cone from p over Δ^1 lies in $U_1 \cup U_2$ and F be the homeomorphism of Δ^2 onto itself that is the identity off of the cone and agrees with f on Δ^1. Then Δ^2 is metrized so that $Fg: D_0^2 \to \Delta^2$ is an isometry. This modification is done for each 1-simplex in $h(B)$. Now the metrics for two 2-simplexes of T with a common edge do agree on the common edge.

CHAPTER Ⅲ THE SCHOENFLIES THEOREM

The methods used in this chapter are special for the plane and most are not known to generalize to all higher dimensions. Some are even false there.

The Schoenflies theorem(Theorem Ⅱ.6.C) has played an important role in Euclidean topology. It shows that if J is any simple closed curve in R^2 whatsoever, then there is a homeomorphism of R^2 onto itself that takes J onto a circle. It is used to discover other useful properties—for example, it is used to show that the Euler characteristic of a 2-sphere is 2. It can be used to show that the cranky neighbor problem has no solution.

Early proofs of the Schoenflies theorem were complicated. Some mathematicians may not have sought proofs since they felt that the result was intuitively obvious and hence near axiomatic. It may have come as a surprise to some to learn that the corresponding theorem in the next dimension is false. Chapter Ⅳ describes wild 2-spheres in R^3.

A proof of the Schoenflies theorem for the plane has been known since 1906. That each simple closed curve separated R^2 into precisely two pieces was already known. Carathéodory is credited with a proof of the Schoenflies theorem from the realm of complex numbers. He used conformal mappings to show that if J is a simple closed curve in R^2, then there is a homeomorphism of the unit disk $|z| \leqslant 1$ onto the union of J and the bounded component of $R^2 - J$. This treatment had the advantage that the homeomorphism preserved right angles for $|z| < 1$, but it had the disadvantage that it gave an existence proof rather than a constructive one and left many students in the dark as to where points went. Taking a unit disk into the disk bounded by J gives one form of the Schoenflies theorem(like Theorem Ⅲ.6.A), but this form is easily translated into others as shown by the proofs of Theorems Ⅲ.6.B and Ⅲ.6.C.

We now have at hand(see Cairns[$\mathbf{C}_1$]and Moise[$\mathbf{M}_{12}$])several good proofs of the Schoenflies theorem. One of the easiest proofs makes use of the PL Schoen-flies theorem. This PL version has merits of its own so we use that approach. We prove the PL Schoenflies theorem in §Ⅲ.1, note some applications of it in §§Ⅲ.2, Ⅲ.3, and Ⅲ.4, and then treat the Schoenflies theorem in Ⅲ.5 and Ⅲ.6.

Ⅲ.1. PL Schoenflies theorem. A *polygon* is a simple closed curve that is sum of a finite number of 1-simplexes. The places where it bends are call *vertices*, and a 1-simplex joining adjacent vertices is called an *edge* or *side*.

THEOREM Ⅲ.1.A(PL SCHOENFLIES THEOREM). *For each polygon P in the pla R^2 there is a finite collection of pushes whose composition sends P to a triangle.*

PROOF. The proof is by induction on the number of sides of P. It is clear the re-





sult holds if P has 3 sides. Let n be an integer larger than 3 such that $3 \leqslant k \leqslant n-1$, Theorem Ⅲ.1. A holds for any polygon with k sides. We then finish the proof by showing that Theorem Ⅲ.1.A holds for any polygon with sides.

Assume P has n sides. It follows from Theorem Ⅲ.1.D(which we assume for now)that there is a nonadjacent pair(say p, q) of vertices of P such that the straight 1-simplex pq intersects P only in its ends. Then P is the union of two PL arcs pxq, pyq. The points x, y are points of the open arcs(pxq), (pyq). They need not be vertices and are used merely to distinguish betwee the three arcs pq, pxq, pyq from p to q. We call $pxq \cup pyq \cup pq$ a θ-curve. The polygons $pxq \cup pq$ *and* $pyq \cup pq$ each has fewer than n edges so Theorem Ⅲ.1.A holds for each of them.

Let h_1, h_2 be finite compositions of pushes in R^2 that take $pxq \cup pq$, $pyq \cup pq$, respectively, to triangles. We do not suppose the images of $p.\ x$, y, q are vertices. Where does $h_1(\text{Int } pyq)$lie? There are two cases according as to whether it lies outside or inside the triangular disk bounder by $r_1(pxq \cup pq)$.

If $h_1(\text{Int } pyq)$does not intersect the 2-simplex bounded by $h_1(pxq \cup pq)$ we find that there is a composition h_3 of three pushes so that $h_3h_1(pxq \cup pq)$ is a triangle with vertices $h_3h_1(p)$, $h_3h_1(q)$, $h_3h_1(x)$. Also, $h_3h_1(\text{Int } pyq)$does not lie in this triangle. As noted in Figure Ⅲ.1.A there is a push h_4 at $h_3h_1(x)$ such that h_4 is fixed on $h_3h_1(pyq)$but $h_4h_3h_1(pxq)$is the 1-simplex $h_3h_1(pq)$. Check that $h_1^{-1}h_3^{-1}h_4h_3h_1$ takes $pxq \cup pyq$ to $\text{pq} \cup \text{pyq}$ and $h_2h_1^{-1}h_3^{-1}h_4h_3h_1$ takes P to a triangle.

If $h_1(\text{Int } pyq)$intersects the 2-simplex bounded by $h_1(pxq \cup pq)$, we interchange the roles of h_1 and h_2 and proceed as before. In this case $h_1(pyq)$lies in the 2-simplex bounded by $h_1(pxq \cup pq)$and $h_1(x)$is accessible from the unbounded components of $h_1(P \cup pq)$. Then $h_2(x)$ does not lie in the triangular disk bounded by $h_2(pyq \cup pq)$ and $h_2(\mathit{Int}(pxq))$does not intersect the disk.

We could have broken our argument for Theorem Ⅲ.1.A with the following halfway theorem. Since we have achieved the whole, we will not belabor the half.

THEORE Ⅲ.1.B. *Suppose Q is a PL θ curve in R^2. If Theorem* Ⅲ.1. *A holds for two of the simple closed curves in θ, it holds for the third.*

We now express the PL Schoenflies theorem in a familiar form which follows from which follows from Theorem Ⅲ.1.A

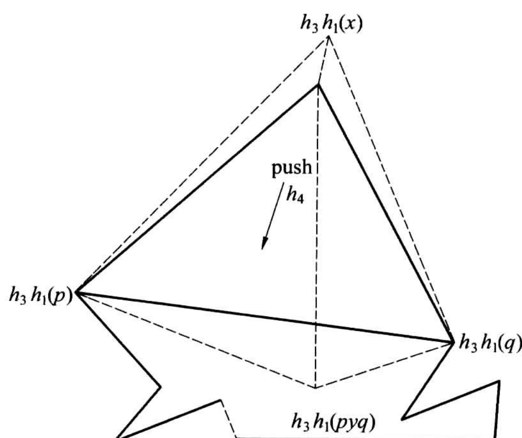


FIGURE Ⅲ.1.A

THEOREM Ⅲ.1.A′(PL SCHOENFLIES THEOREM). *For each pair of polygons P_1, P_2 in R^2 there is a PL homeomorphism with compact support of R^2 onto itself that takes P_1 to P_2.*

There are generalizations of the above theorem saying how we can keep the required PL homeomorphism fixed on certain sets or make the PL homeomorphism agree with certain preassigned PL homeomorphisms of P_1 onto P_2. However, we will not pursue these.

By the PL Schoenflies theorem we can speak not only of the interior and exterior of a triangular polygon but also of any polygon P in R^2. Int P and Ext P are the bounded and unbounded components respectively of $R^2 - P$.

THEOREM Ⅲ.1.C. *For each pair of polygons P_1, P_2 in R^2, each homeomorphism h of P_1 onto P_2 can be extended to a homeomorphism of $R_1 \cup$ Int P_1 onto $P_2 \cup$ Int P_2.*

PROOF. Let g_i be a homeomophism of $P_i \cup$ Int P_i onto a triangular disk D as guaranteed by THEOREM Ⅲ.1.A. Select a point p_0 of Int D and let $h': D \to D$ be the homeomorphism such that if $y \in$ Bd D, h' takes the 1-simplex $p_0 y$ linearly onto the 1-simplex from p_0 to $g_2 h g_1^{-1}(y)$. Then $g_2^{-1} h' g_1$ is a suitable extension of h.

THEOREM Ⅲ.1.D. *If P is a polygon with more than three vertices in R^2, some pair of vertices of P can be joined by a 1-simplex that intersects P at only these two points.*

PROOF. Let v_1, v_2, v_3 be three consecutive vertices of P. If $v_1 v_3$ does not ser for the required 1-simplex, $(v_1 v_3)$ intersects P. If Int $v_1 v_2 v_3$ contains a vertex P, let v be the farthest such vertex from the line through v_1, v_3 and note that v will serve. If$(v_1 v_3)$



intersects P but no vertex of P lies in Int $v_1v_2v_3$, then(v_1v) contains a vertex v of P and v_2v will serve.

ALTERNATE PROOF. Those acquainted with convex sets and lines of support may prefer the following proof. If P bounds a convex disk, any nonadjacent pair of vertices will serve. If it does not bound a convex disk, some line of support intersects P in a nonconnected set. This line of support contains a 1-simplex verifying the theorem.

Ⅲ. 2. Triangulating PL disks. This section gives some applications of the PSchoenflies theorem.

A PL *disk* is the homeomorphic PL image of a 2-simplex. A polygon P in R^2 bounds a disk D and we write Bd $D = P$. Sometimes a PL disk is called *a polygonal disk*. If$(X, T)$is a complex, we say that X is PL. Here the symbol PL indicates polygonal or polyhedral. Recall from § Ⅱ.2 that it meant piecewise linear when referring to a map.

THEOREM Ⅲ. 2. A. *If Q is a PL θ-curve in R^2, the three disks D_1, D_2, D_3 bounded by the three polygons in Q can be ordered D_1, D_2, D_3 so that $D_3 = D_1 \cup D_2$ and $D_1 \cap D_2$ is an arc in Q.*

We plan to get an efficient triangulation of a polygonal disk D in R^2. To this end we consider certain spanning arcs of D. An arc ab in D *spans* D if it intersects Bd D only in the arc's ends. A spanning arc ab is a *proper spanning segment* if it is straight and its ends are vertices of Bd D.

THEOREM Ⅲ.2.B. *Each PL disk D in R^2 with more than 3 sides has a proper spanning segment.*

PROOF. Let v_1, v_2, v_3 be three consecutive vertices of Bd D. The proof breaks into two cases according as to whether or not the part of Int $v_1v_2v_3$ near v_2 lies in D.

If it lies in D, we use the technique of the proof of Theorem Ⅲ.1.D to obtain a proper spanning segment.

If it lies in D, we use the technique of the proof of Theorem Ⅲ.1.D to obtain a proper spanning segment.

If it does not lie in D, we extend a ray from v_2 in a direction that starts the ray into Int D. The open ray first hits P at a point x. If x is a vertex of Bd D, v_2x is a proper spanning segment. If it is not a vertex, x lies on a side ab of Bd D—pick a to be the one of these such that neigher v_1 or v_3 lies in the 2-simplex axv_2. See Figure Ⅲ.2.B. Suppose v_2a is not a proper spanning segment. Let y be the first point of xa in the order from x to a such that v_2y intersects Bd D in more than two points. A subarc of ay is a spanning segment from v_2.

Note that we have not shown that there is a proper spanning segment from v_2. However there is if v_1v_3 is not a proper spanning segment.

Repeated applications of Theorem Ⅲ.2.B gives the following result.

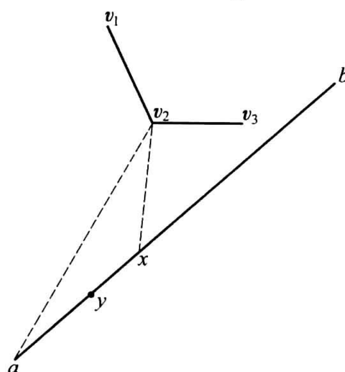


FIGURE Ⅲ.2.B

THEOREM Ⅲ.2.C. *If P is a polygon in R^2, then there is a triangulation T of the polygonal disk bounded by P such that each 0-simplex of T is a vertex of P and each 1-simplex of T is either an edge of P or a proper spanning segment of P.*

COROLLARY Ⅲ.2.D. *If P is a polygon in R^2, then there is a triangulation T of the convex hull of p such that each vertex of T is a vertex of P and each edge of P is a 1-simplex of T.*

Ⅲ.3. Skew curves. We give some further applications of the PL Schoenflies theorem. A *finite graph* is the union of the finite number of arcs such that if two of them intersect, their intersection is an end point of each. We shall be concerned with restrictions on the embeddability of certain finite graphs in R^2.

THEOREM Ⅲ.3.A. *Suppose a_1, b_1, a_2, b_2 are points as ordered on a polygon P in R^2 and a_1a_2, b_1b_2 are disjoint PL arcs which intersect P only at their ends. Then (a_1a_2), (b_1b_2) lie in different components of R^2—P.*

PROOF. Apply Theorem Ⅱ.2.A to $P \cup a_1a_2$ as shown in Figure Ⅲ.3.A. Note that (b_1b_2) lies in one of Int D_1, Int D_2, Ext D_3.

If a_1a_2 is the middle arc $D_1 \cap D_2$, b_1 and b_2 are on outside arcs and $(b_1b_2) \subset$ Ext D_3. In this case P separates (a_1a_2) from (b_1b_2).

If a_1a_2 is the middle arc $D_1 \cap D_2$, b_1 and b_2 are on outside arcs and $(b_1b_2) \subset$ Ext D_3. In this case P separates (a_1a_2) form (b_1b_2).

If a_1a_2 is the top arc, $(b_1b_2) \subset$ Int D_2 and P separates (a_1a_2) from (b_1b_2).

If a_1a_2 is the bottom arc, $(b_1b_2) \subset$ Int D_1 and P separates (a_1a_2) from (b_1b_2).

We shall show that certain finite graphs cannot be embedded in R^2. In view of our reliance ont the PL Schoenflies theorem we show that if certain graphs could be embedded, topologically equivalent PL graphs could be embedded also.

THEOREM Ⅲ.3.B. *If p, q are two points of a connected open set U of R^2, there is a PL arc in U from p to q.*

THEOREM Ⅲ.3.C. *For each finite graph G in R^2 and each $\varepsilon > 0$ there is a homeomorphishm $h: G \to R^2$ such that $d(h, \mathrm{Id}) < \varepsilon$ and $h(G)$ is PL.*

PROOF. Subdivide arcs making up G so that G is the union of very small arcs A_1, $A_2, \cdots, A_n$ such that if two of the A's intersect, their intersection is an end of each. Put very very small round disks about the ends of the A's and replace each A by the union of a radial arc in each of the two disks about the ends and a PL arc near A_i joining the radial ends.

A *skew curve of type* 1 is a set homeomorphic to the 1-complex shown in Figure Ⅲ.3.D. This 1-complex has six vertices a_1, a_2, a_3, b_1, b_2, b_3 and nine arcs $a_1 b_1$, $a_1 b_2, \cdots, a_3 b_3$ as shown. The arcs are not assumed to be straight. The curve is sometimes called the cranky neighbor graph (or the utility example) because of the question as to whether or not three neighborhoods a_1, a_2, a_3 can run utility lines to three stations b_1, b_2, b_3 so that their utility lines will not cross.

THEOREM. Ⅲ.3.D. *No skew curve of type* 1 *lies in the plane.*

PROOF. In view of Theorem Ⅲ.3.C we suppose that the skew curve of type 1 under consideration is PL. Suppose $a_1 b_1 a_2 b_2$ is the P of Theorem Ⅲ.3.A and $(a_1 b_3 a_2)$, $(b_1 a_3 b_2)$ are the open arcs $(a_1 a_2)$, $(b_1 b_2)$ of that theorem. If the skew curve lies in R^2, $a_3 b_3$ shows that P does not separate $(a_1 b_3 a_2)$ from $(b_1 a_3 b_3)$. This viloates Theorem Ⅲ.3.A.

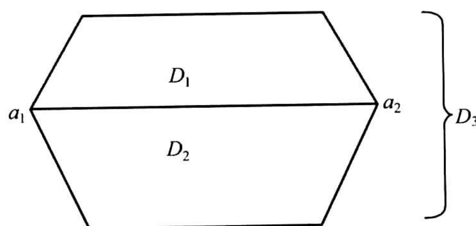


FIGURE Ⅲ.3.A

A *skew curve of type* 2 is a set homeomorphic to the complete graph on five vertices as shown in Figure Ⅲ.3.E. Kuratowski showed that a finite graph could be embedded in the plane if it did not contain either a skew curve of type 1 or one of type 2.

THEOREM Ⅲ.3.E. *No skew curve of type* 2 *lies in the plane.*

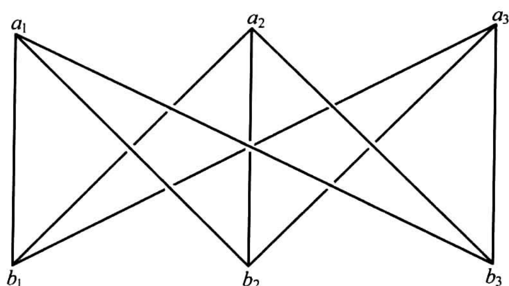


FIGURE Ⅲ.3.D

PROOF. Again we assume the graph under consideration in PL and assume it lies in the plane. We let $a_1a_2a_3a_4a_5$ be the P of Theorem Ⅲ.3.A and denote the components of R^2-P by U and V where$(a_2a_5)\subset U$. It follows from Theorem Ⅲ.3.A that $(a_1a_3)$and(a_1a_4)lie in V while(a_2a_4), (a_3a_5)lie in U. This contradicts Theorem Ⅲ.3.A which states that$(a_2a_4)$and(a_3a_5)lie in different components of R^2-P. The contradiction arose from the faulty assumption that some skew curve of type 2 lies in R^2.

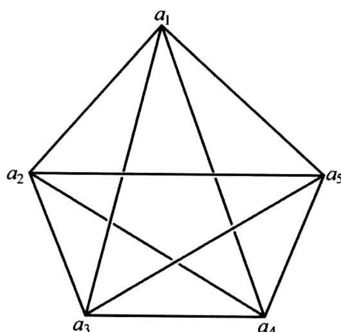


FIGURE Ⅲ.3.E

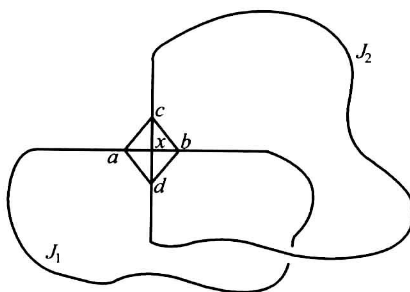
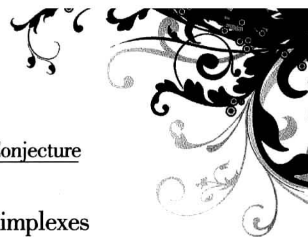


FIGURE Ⅲ.3.F

Ⅲ.4. No arc separaies R^2. It is sometimes useful to partition a set. If T is a





triangulation of the boundary of a tetrahedron, the collection of interiors of 2-simplexes of T is a partitioning of the boundary. A *partitioning* of a set X is a finite collection of mutually disjoint open subsets of X whose union is dense in X.

A partitioning is a *brick partitioning* if each of its elements is the interior of its closure, is uniformly locally connected, and if the interior of the closure of the union of each two elements is also uniformly locally connected. Here *uniformly locally connected* means 0-ulc as defined in § 5 of Chapter **VIII**. One can get a brick partitioning of the boundary of a tetrahedron by letting T be a triangulation of the boundary, T' be the barycentric subdivision of T, and letting the elements of the partitioning be the open stars in T' of the vertices of T.

Bing used brick partitionings in[$\mathbf{B}_6$] to show that each continuous curve has a convex metric. He used them in[$\mathbf{B}_3$] to improve the Kline sphere characterization. We need only a mild form of it in the proof of Theorem **III.4.A**, and there we may use as elements of the partitionintgs the interiors of rectangles stacked so that the interior of the cloure of the union of any two of them is uniformly locally connected. The boundary of the closure of the union of any subset of them is the union of a finite number of mutually disjoint polygons.

There are many proofs of the result that no arc separates R^2. The one we use here makes a pleasant application of the PL Schoenflies theorem. A variation of the methods used here can be used to get related results in higher dimensions. See Chapter VIII for some of these extended results.

THEOREM III.4.A(JANISZEWSKI). *Suppose that in* R^2

A, B *are two closed sets*,

A *is compact*,

p, q *are two points of* $R^2 - (A \cup B)$,

paq is a polygonal arc from p *to* q *in* $R^2 - B$,

pbq is a polygonal arc from p *to* q *in* $R^2 - A$, *and*

$A \cap B$ *is connected or null.*

Then there is a polygonal arc pxq from p to q such that

$$pxq \subset R^2 - (A \cup B)$$

PROOF. We suppose that the theorem is false, that $paq \cap pbq$ is finite, and that r is the first point of $paq \cap pbq$ in the order from p to q on paq such that p and r belong to different components of $R^2 - (A \cup B)$. For convenience we suppose that $r = q$, that $paq \cup pbq$ bounds a rectangular disk D as shown in Figure III.4.A, and that $B \cap \mathrm{qa} = \emptyset$.

First we consider the case where $A \cap B \cap D = \emptyset$. Let ε be a positive number less than $d((A \cup pa), (B \cup bq))$ and $\{D_i\}$ be a brick partitioning of D into rectangular disks as shown in Figure Ⅲ.4.A so that each D_i is of diameter less than ε and each component of the boundary of the union of a finite number of the D_i's is a polygon. Let X be the union of the D_i's intersecting $A \cup pa$ and P be the component of Bd X containing pa. Then $(P - (pa)) \cup qa$ contains an arc in $D - (A \cup B)$ joining p and q which shows the assumption that the theorem is false is wrong.

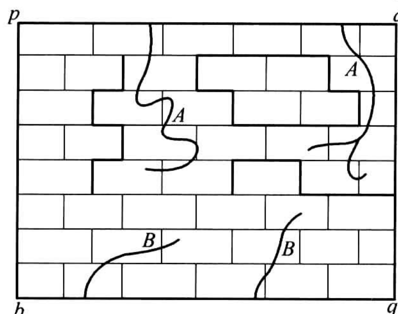


FIGURE Ⅲ.4.A

The treatment in case $A \cap B$ misses $R^2 - D$ is similar except that we partition $R^2 - \text{Int } D$ instead of D. There are infinitely many D_i's but only a finite number intersect the compact set A.

THEOREM Ⅲ.4.B. *No arc (topological) in* R^2 *separates* R^2.

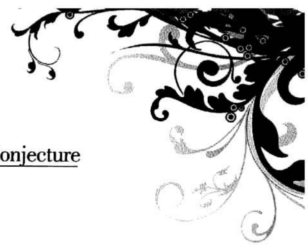
PROOF. Suppose an arc ab in R^2 separates p from q. Let $a = a_1, a_2, \cdots, a_n = b$ be points on ab so close together than no a_ia_{i+1} separates p from q. Iterated use of Theorem Ⅲ.4.A shows that $a_1a_3, a_1a_4, \cdots, a_1a_n = ab$ does not either.

Ⅲ.5. Jordan-Brouwer theorem. The Jordan curve theorem says that each simple closed curve J in R^2 separates R^2 and each component of $R^2 - J$ has J for a boundary. The Jordan-Brouwer theorem puts on the additional conclusion that $R^2 - J$ has exactly two components. The Jordan-Brouwer theorem plays an important role in plane topology. R.L. Moore[$\mathbf{M}_{14}$] used it as one of the axioms in one of this axiomatic treatments of plane topology.

The Jordan-Brouwer theorem is weaker than the Schoenflies theorem but unlike it, remains true for $(n-1)$-spheres in R^n. We use a proof based on applications of the PL Schoenflies theorem because the proof is easy. Later we will shift to a more complicated proof for higher dimensions.

A simple closed curve in R^2 is a *semipolygon* if it contains a straight arc. The open straight are is called a *flat spot*.





THEOREM III.5.A. *A semipolygon J in R^2 separates R^2.*

PROOF. Let axb be a straight arc in J and cxd be a short straight arc intersecting J only at x. Then J separates c from d or else there is a simple closed curve J_2 containing cxd, intersecting J only at $\{x\}$, and violating Theorem Ⅲ.3.F.

The following is an important application of Theorem III.4.B.

THEOREM III.5.B. *If S is a semipolygon in R^2 and U is component of $R^2 - S$, then each point S is a limit point of U.*

A set X is accessible from a set Y at a point $p \in X$ if there is an arc pq in $X \cup Y$ such that pq lies in $Y \cup \{p\}$. If pq is straight, we say p is accessible from Y along a straight arc. The following is an important application of Theorem III.4.B.

THEOREM III.5.C. *If J is a simple closed curve in R^2 and U is a component of $R^2 - J$, then there is a dense sex X of points of J such that if $p \in X$, J is accessible at p from U along a straight arc.*

PROOF. We show that for each $q \in J$ and each neighborhood N of q there is a point $p \in N \cap J$ such that p is accessible from U along a straight arc. Let ab be an arc in $J - \{q\}$ which contains J-N. It follows from Theorem Ⅲ.4.B that there is a PL arc cq in $R^2 - ab$ from a point c of U to q. The first point of J on cq in order from c is a suitable p.

Theorem Ⅲ.5.C implies the following.

THEOREM Ⅲ.5.D. *If J is a simple closed curve in R^2 and U is a component of $R^2 - J$, then each point of J is a limit point of U.*

THEOREM Ⅲ.5.E. *Each simple closed curve J in R^2 separates R^2.*

PROOF. Let pq be a straight arc that intersects J only at its ends. Denote the two arcs in J from p to q by paq, pbq. Let D be a round disk about a which misses $pbq \cup pq$. It follows from Theorems Ⅲ.5.A and Ⅲ.5.C that there are points r, s in D which belong to different components of $R^2 - (paq \cup pq)$. We shall show that J separates r from s.

Assume J does not separate r from s and rcs is a PL arc in $R^2 - J$ from r to s. We adjust rcs to reduce its intersection with pq. With no loss of generality we suppose $c \in pq$ and rcs intersects pq only at a flat spot of rcs. Let rs be a straight arc in D and J_2 be a simple closed curve in $rs \cup rcs$ that contains c.

Then J_2 separates p from q because of the crossing of pq and rcs. However it does not separate them because of pbq. The assumption that J did not separate r from s led to this contradiction.

THEOREM Ⅲ.5.F. *If J is a simple closed curve in R^2, $R^2 - J$ does not have more*

than two components.

PROOF. Suppose $R^2 - J$ has 3 components U_1, U_2, U_3. Let $a_1, b_1, a_2, b_3, a_3, b_2, a_1$ be six points on J ordered as listed. It follows from Theorems Ⅲ.3.B and Ⅲ.5.C that there is an open arc from a point of J near a_1 to a point of J near b_3 in U_1, an open arc from a point of J near b_1 to a one near a_3 in U_2, and an open arc from a point of J near a_2 to one near b_2 in U_3. This contradicts Theorem Ⅲ.3.D.

We combine Theorems Ⅲ.5.C, Ⅲ.5.D, Ⅲ.5.E, and Ⅲ.5.F to get the following.

THEOREM Ⅲ.5.G (JORDAN-BROUWER THEOREM). *If J is a simple closed curve in R^2, then $R^2 - J$ has exactly two components and is the boundary of each.*

After a treatment of the Jordan-Brouwer theorem we are in position to speak of the interior of any simple closed curve whatsoever in R^2. Theorem Ⅲ.2.A can be restated as

THEOREM Ⅲ.5.H. *If Q is a θ-curve in R^2, the three simple closed curves in Q can be ordered J_1, J_2, J_3 so that $J_3 \cup \text{Int } J_3 = (J_1 \cup \text{Int } J_1) \cup (J_2 \cup \text{Int } J_2)$ and $(J_1 \cup \text{Int } J_1) \cap (J_2 \cup \text{Int } J_2)$ is an arc in Q.*

Ⅲ.6. Schoenflies theorem. We now come to the crux of the proof of the Schoenflies theorem. In a basic preliminary step we extend the homeomorphism h of a simple closed curve J in R^2 onto a triangular polygon to a homeomorphism of $J \cup \text{Int } J$ onto the triangular disk bounder by the polygon. The proof uses infinitely many steps. Figure Ⅲ.6.A.a suggests a start. Theorems Ⅲ.6.A, Ⅲ.6.B, Ⅲ.6.C are all versions of the Schoenflies theorem.

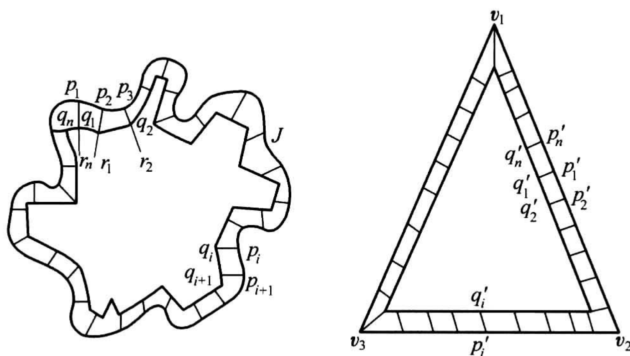
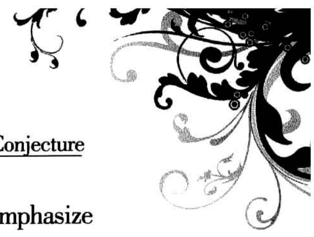


FIGURE Ⅲ.6.A.a

THEOREM Ⅲ.6.A. *Suppose in R^2 that J is a simple closed curve and D is a triangular disk. Any homeomorphism h of J onto* Bd D *can be extended to take $J \cup \text{Int } J$ onto D.*



PROOF. We use a sequential description of the extension of h and to emphasize that we will be making repeated uses of the same technique, we let $\varepsilon_1, \varepsilon_2, \ldots$ be a sequence of positive numbers converging to 0. The harmonic sequence $1, \frac{1}{2}, \frac{1}{3} \ldots$ would do, but the choice is unimportant.

Theorem Ⅲ.5.C allows us to pick points $p_1, p_2, \ldots, p_n$ cyclically orderd on J so abundant and so close together that the arcs $p_i p_{i+1}$ on J and their images on BD D are small(have diameters less than $\varepsilon_1/3$)and the p_i's are accessible from Int J by straight arcs. (We treat p_{n+1} as p_1.)

Let $p_i r_i$ be mutually disjoint short(diameters less than $\varepsilon_1/3$)straight arcs which lie. except for their end p_i, in Int J. Designate $h(p)$ on Bd D by p'_1 and let $p'_i q'_i$ be mutually disjoint short(diameters less than $\varepsilon_1/3$)straight arcs which lie except for p'_i in Int D.

If follows from repeated applications of Theorem Ⅲ.6.D(which we will prove later)that there is a polygon $q_1 q_2 \cdots q_n$ in Int J(the polygon may have more than n vertices)and straight arcs $p_1 q_1, p_2 q_2, \cdots, p_n q_n$ as shown in Figure Ⅲ.6.A.a Here each $p_i q_i \cup q_i q_{i+1} \cup q_{i+l} p_{i+1}$ lies in an $\varepsilon_1/3$ neighborhood of $p_i p_{i+1}$. The polygon $q_1 q_2 \cdots q_n$ may be constructed as follows. Apply Theorem Ⅲ.6.D directly to get an arc $s_1 s_2$ in U such that $s_1 s_2$ is close to $p_1 p_2$ and intersects $p_i r_i$ only at $s_i (i = 1,2)$. We suppose $p_1 s_1 \cup s_1 s_2 \cup s_2 p_2$ misses superfluous $p_j r_j$'s. Let O_2 be a neighborhood of $p_2 p_3$ that misses $s_1 s_2 \cup s_2 r_2$ and superfluous $p_j r_j$'s and apply Theorem Ⅲ.6.D to get an arc $q_2 s_3$ in $O_2 \cap U$ irreducible from $p_2 r_2$ to $p_3 r_3$. Note that q_2 is between s_2 and p_2. We continue building $q_3 s_4, q_4 s_5, \cdots, q_n q_1$. The $q_1 q_2 \cdots q_n$ is a union of $s_1 s_2$, ($s_2 p_2$ of $p_2 r_2$), $q_2 s_2$, ($s_3 q_3$ of $p_3 r_3$), $\cdots$, $q_n q_1$, ($q_1 s_1$ of $p_1 r_1$).

We now extend h to take the $p_i q_i$'s to the $p'_i q'_i$'s and the polygon $q_1 q_2 \cdots q_n$ to $q'_1 q'_2 \cdots q'_n$. It is further extended to take the PL disk in Int J bounded by $q_1 q_2 \cdots q_n$ onto the disk in Int D bounded by $p'_1 q'_2 \cdots q'_n$. This last extension is allowed by Theorem Ⅲ.1.C.

We have now extended h to define it on all of $J \cup$ Int J except on the interiors of some small simple closed curves. The $\varepsilon_1/3$ restriction insures that these simple closed curves and their images in D have diameters less than ε_1. A typical simple closed curve is shown in **Figure Ⅲ.6.A.**a as $p_i q_i \cup (q_i q_{i+1}$ of $q_1 q_2 \cdots q_n) \cup q_{i+1} p_{i+1} \cup (p_{i+1} p_i$ of J)and in **Figure Ⅲ.6.A.**b as $J_i = p_i r_i q_i \cup q_i b_i q_{i+1} \cup q_{i+1} r_{i+1} p_{i+1} \cup p_{i+1} a_i p_i$.

Let r_i be a point of($p_i q_i$)within $\varepsilon_2/3$ of $p_i i$. It follows from Theorems **Ⅲ.6.D** and **Ⅲ.5.C**, and a procedure previously used that there is a PL arc $r_i c_i r_{i+1}$ in $\{r_i\} \cup$

$\{r_{i+1}\} \cup \text{Int } J_i$ as shown in Figure Ⅲ.6.A.b and straight arcs joining this arc to $p_i a_i p_{i+1}$ as shown so that each of the small upper simple closed curves in Figure Ⅲ.6.A.b has diameter less than ε_2.

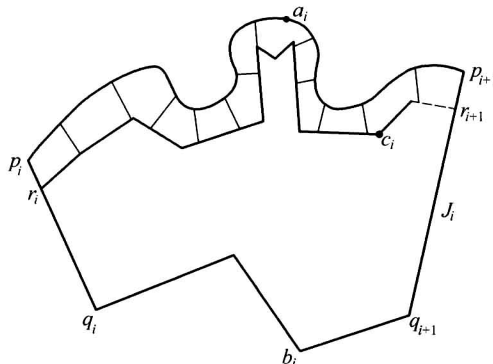


FIGURE Ⅲ.6.A.b

We build such a PL configuration in the PL disk in D bounded by the image of J_i and extend h to take the lower PL disk in **Figure Ⅲ.6.A.**b(as well as the straight arcs shown joining $r_i c_i r_{i+1}$ to $p_i a_i p_{i+1}$) into this PL disk. This is done so that the mage of each of the small upper simple closed curves in **Figure Ⅲ.6.A.**b has a diameter of less than ε_2.

The extension is made for each J_i and we move onto the next step which is a repetition of the last except that we use ε_3 instead of ε_2. The extension is continued and the limit is the required extension. If $p \in \text{Int } J$, h is continuous at p since some neighborhood of p eventually misses the small disks at a far-out stage on which h has not been precisely defined. If $p \in J$, h is continuous at J, since the far-out disks are small and are sent by h(enen imprecisely at this stage)into disks bounded by small simple closed curves.

With Theorem **Ⅲ.6.A** we can lift a restriction from Theorem **Ⅲ.1.C.**

THEOREM **Ⅲ.6.B** *For each pair of simple closed curves J_1, J_2 in R^2, each homeomorphism ϕ of J_1 onto J_2 can be extended to a homeomorphism h of $j_1 \cup$ Int J_1 onto $J_2 \cup$* Int J_2.

PROOF. It follows from Theorem **Ⅲ.6.A** that there is a homeomorphism of $J_1 \cup$ Int J_1 onto a triangular disk and a homeomorphism g_2 of $J_2 \cup \text{Int } J_2$ onto D that agrees with $g_1 \phi^{-1}$ on J_2. Then $g_2^{-1} g_1 = h$ satisfies the requirements.

THEOREM **Ⅲ.6.C.** *For each pair of simple closed curves J_1, J_2 in R^2 there is a homeomorphism h of R^2 onto itself such that $h(J_1) = J_2$ and h has compact support.*

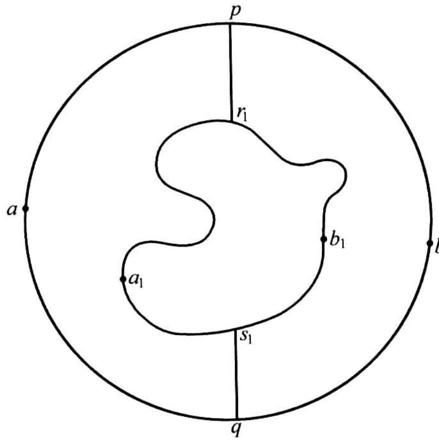


FIGURE Ⅲ.6.C

THEOREM Ⅲ.6.D. *Suppose p_1p_2 is an arc of simple closed curve J, U is the bounded component of $R^2 - J$, and p_1r_1, p_2r_2 are straight arcs which lie except for p_1, p_2 in U. Then for each neighborhood O of p_1p_2 there is an arc s_1s_2 in $O \cap U$ such that $p_ir_i \cap s_1s_2$ for $i = 1,2$.*

PROOF. Suppose $p_1r_1cr_2p_2$ is a PL arc containing $p_1r_1 \cup p_2r_2$ which lies except for its ends in U. Use a brick partitioning (such as shown in Figure Ⅲ.4.A) of part of R^2 to get a connected open set N such that $r_1cr_2 \cap N = \emptyset$, $p_1p_2 \subset N \subset \overline{N} \subset 0$, each component of Bd N is a polygon, while Bd N and $p_1r_1cr_2p_2$ have triangulation that make them in general position. We suppose some point $b \in J$ lies in the unbounded component of $R^2 - \overline{N}$. By cutting slices from N along parts of $p_1r_1cr_2p_2$, we reduce N until Bd $Np_1r_1cr_2p_2 \cap$ Bd N consists of just two points s_1, s_2 of $p_1r_1p_2r_2$, respectively, as shown in Figure Ⅲ.6.D.

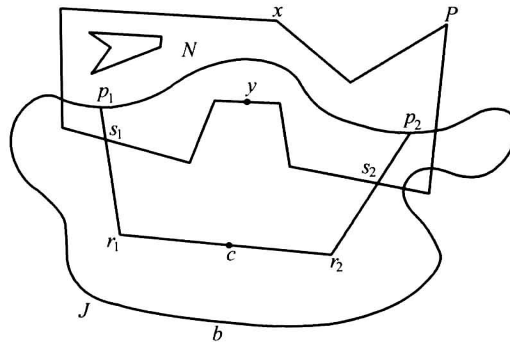


FIGURE Ⅲ.6.D

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

The arc s_1ys_2 of Figure Ⅲ.6.D satisfies Theorem Ⅲ.6.D so we finish the proof of Theorem Ⅲ.6.D by showing how to get s_1ys_2. Note that s_1, s_2 belong to the same boundary component P of Bd N since Bd N crosses the simple closed curve $K = p_1q_1 \cup pj_1r_1cr_2p_2$ only at s_1 and s_2. Suppose s_1 and s_2 separate x from y on P where $y \in$ Int K. Then s_1ys_2 lies except for its ends in Int K. Since Int K does not intersect J, $s_1ys_2 \subset U$.

CHAPTER Ⅳ　WILD 2-SPHERES

It is natural to try to generalize to higher dimension interesting concepts and theorems from the plane. One of those that does not generalize is the Schoenflies theorem. There is a 2-sphere in R^3 such that there is no homeomorphism of R^3 onto itself taking the 2-sphere onto the boundary of a round ball. In this chapter we give many such examples.

Perhaps more details are included in this chapter than the reader will want to assimilate before moving to later chapters. Some feel that one should master all previous material moving to later chapters. Some feel that one should master all previous material before trying to forge ahead but that is not the way that Chapter Ⅳ was constructed. The material on the Alexander horned sphere and Antoine's necklace is basic. That on involutions, Fox-Artin spheres, and hooked rugs is for breadth and is not required for work in future chapters. A prudent plan of action might be to become familiar with some of the basic examples and then return to study more complicated ones as one gets better acquainted with the complexities of R^3. Examples of unusual spheres were given by Alexander [**A$_3$**, **A$_4$**], Alford [**A$_6$**, **A$_7$**], Ball [**B$_1$**], Bing [**B$_5$**, **B$_{13}$**, **B$_{31}$**], Cannon [**C$_6$**], Càsler [**C$_7$**], Daverman and Eaton [**DE$_1$**, **DE$_2$**], Eaton [**E$_2$**], Fort [**F$_2$**], Fox and Artin [**FA**], Hempel [**H$_7$**], Martin [**M$_1$**, **M$_2$**], Rolfsen [**R$_1$**], Rosen [**R$_2$**].

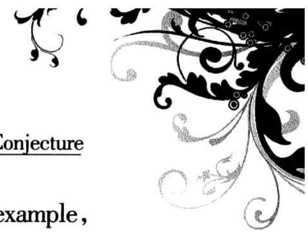
Ⅳ.1. Tame and wild 2-spheres. An n-sphere is an object homeomorphic with the boundary of an $(n+1)$-ball. The *canonical n-sphere* is the set of points in R^{n+1} satisfying the equation

$$x_1^2 + x_2^2 + \cdots + x_{n+1}^2 = 1$$

We use S^n to denote this canonical n-sphere. Any object topologically equivalent to this canonical n-sphere is called an n-sphere and on occasions it is convenient to denote one of these other *n-spheres* by S^n. If B^{n+1} is an $(n+1)$ - *ball* and $=_t$ means is *topologically equivalent to*, then

$$\text{Bd } B^{n+1} =_t S^n$$





The above definition has more intuitive meaning in lower dimensions. For example, one sees that

$$S^0 =_t \text{pair of points}$$

$$S^1 =_t \text{circle} =_t \text{simple closed curve}$$

$$S^2 =_t \text{boundary of a round ball}$$

Also, the boundaries of ellipsoids, rectangular solids, and tetrahedra are examples of 2-spheres. The boundary of a rough rock might be a more exotic example. It is more difficult to visualize examples of higher-dimensional spheres. There are not such readily available examples of n-spheres for $n = 3$ in Euclidean 3-space R^3, the space in which we presume we live.

A 2-sphere S^2 in R^3 is called *tame* if there is a homeomorphism of R^3 onto itself that takes S^2 onto the boundary of a round ball. If there is no such homeomorphism, S^2 is called *wild*.

If S^2 is a 2-sphere in R^3, the union of S^2 and its bounded complementary domain is called a *crumpled cube*. (Think of "the cow with the crumpled horn".) If S^2 is wild, there is no assurance that the associated crumpled cube is homeomorphic with a 3-ball. If S^2 lies ini S^3, rather than in R^3, the union of S^2 and either of its complementary domains is called a *crumpled cube*. Any set homeomorphic with a crupled cube is called a crumpled cube irrespective of where it is embedded.

If(X, T) is a complex we may call X a *geometric complex* to emphasize that we are not associating any particular triangulation with it. A subset Y of a geometric complex X is called a *polyhedron* if there is a triangulation T of X(compatible with X's linear structure)such that Y is the union of simplexes of T. In general, a subset W of a geometric complex X is called *tame* if there is a homeomorphism of X onto itself that takes W onto a polyhedron. Recall that we have defined simplexes to be closed and therefore a tame subset of X is closed in X.

IV.2. 3-spheres. A 3-sphere may be more difficult to visualize than a lower-dimensional sphere. The 3-sphere is of primary concern in the study of 3-manifolds. We give special attention to S^3 and describe four models of it.

IV.2.A. S^3 *as the boundary of a* 4-*simplex*. If one wants to consider S^3 with a PL structure, an approach is to regard S^3 as the boundary of a 4-simplex. Looking at S^3 with the triangulation engendered by the 4-simplex, one would see S^3 as the union of five tetrahedra. A disadvantage of such an approach is that we have used a complicated object(4-simplex)to define a simpler one.

IV.2.B. S^3 *as the union of balls*. It is sometimes convenient to regard S^3 as the u-

nion of two 3-balls. Before describing this model, we first consider some lower-dimensional aspects of such a model.

One may decompose a circle into two semicircles. Hence, if one takes two disjoint arcs and uses a homeomorphism to sew the end points of one arc onto the end points of the other, one has an example of a 1-sphere. One can get a model of a 2-sphere by considering two disks D_1, D_2 and joining them along their boundaries. If h is the homeomorphism of Bd D_1 onto Bd D_2 such that each point p of Bd D_1 is fitted to the point $h(p)$ of BD d_2, we may write

$$S^2 = {}_tD_1 \cup_h D_2$$

Similarly, one may obtain a model of S^3 by sewing two balls B_1^3, B_2^3 together with a homeomorphism h of Bd B_1^3 onto Bd B_2^3. The sewing is abstract rather than physical since it cannot be realized in R^3. See Figure IV.2.B. We write

$$S^3 = {}_tB_1^3 \cup_h B_2^3$$

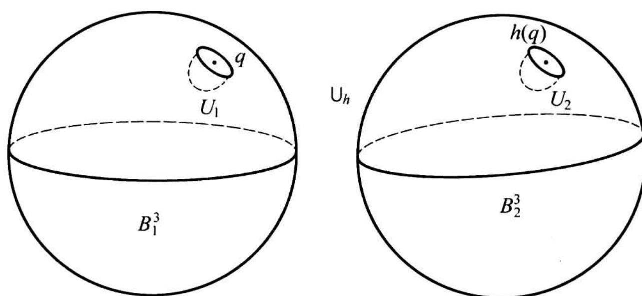


FIGURE IV.2.B

The topology of $B_1^3 \cup_h B_2^3$ is as follows. If $P \in$ Int B_i^3, then any open set in Int B_i^3 containing p is a neighborhood of p in $B_1^3 \cup_h B_2^3$. (We define a *neighborhood of p* as an open set containing p.) If $q \in BD\ B_1^3$, then $\{q\} \cup \{h(q)\}$ is considered a point in $B_1^3 \cup_h B_2^3$. To get a neighborhood of such a $\{q\} \cup \{h(q)\}$, one considers an open set U_1 in B_1^3 containing q and an open set U_2 in B_2^3 containing $h(q)$ such that $h(U_1 \cap \text{Bd } B_1^3) = U_2 \cap \text{Bd } B_2^3$. The associated neighborhood of $\{q\} \cup \{h(q)\}$ is

$$U_1 \cup_{h|(U_1 \cap \text{Bd } B_1^3)} U_2$$

See Figure IV.2.B.

IV.2.C. S^3 *as the one point compactification of* R^3. Before describing S^3 as the one-point compactification of R^3, we consider a 2-dimensional version of this model.

Suppose S^2 is a round 2-sphere tangent to the plane R^2 at a point p. Let q be the point of S^2 diametrically opposite to p. Consider the projection from q of $S^2 - \{q\}$ onto

R^2. If $x \in S^2 - \{q\}$, its image is the point where the line joining q and x passes through R^2. Hence

$$S^2 - \{q\} =_t R^2$$

This suggests that the one point compactification of R^2 is a model for S^2—the point of compactification added to R^2 corresponds to q.

The points of the 1-point compactification of R^3 are the points of R^3 together with an ideal point which we call the point at infinity. Let us use p_∞ to denote this ideal abstract point. The canonical neighborhoods in this enlarged R^3 are of two sorts: (i) the interior of a ball in R^3; (ii) the union of $\{p_\infty\}$ and the complement in R^3 of a closed 3-ball with center at the origin—the larger the 3-ball, the smaller the neighborhood of p_∞.

Actually, it is easy to see that this model of S^3 is equivalent to the one given in the preceding § IV.2.B. Suppose B_1^3 is the ball in R^3 with its center at the origin and radius 1. Then $\{p_\infty\} \cup (R^3 - \text{Int } B_1^3)$ may be considered as a topologica ball B_2^3 with center at p_∞. If r is a ray reaching out from the origin and r' is the part of this ray outside Int B_1^3, then $\{p_\infty\} \cup r'$ can be considered a radius of B_2^3. Then the 1-point compactification of R^3 is the union of B_1^3 and B_2^3 sewn together with the identity map of Bd B_1^3 onto Bd B_2^3. See Figure IV.2.C.

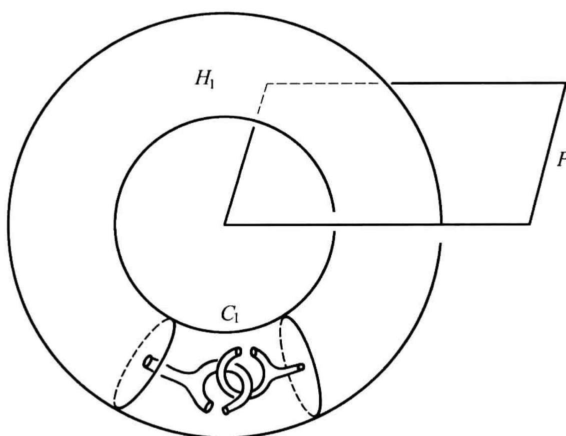


FIGURE IV.2.C

IV.2.D. S^3 *as the union of two solid tori*. Consider two solid tori T_1^3, T_2^3 linked as shown in Figure IV.2.D. Each of these tori is topologically like a doughnut (that is, like $D^2 \times S^1$). Can the two tori be blown up until their boundaries fit? If this could be done, their union would be a model of S^3. There is a bit of a problem in getting one of them to swallow the point at infintiy. Perhaps one could visualize this expansion better

after considering a more precise abstract approach. We describe one below.

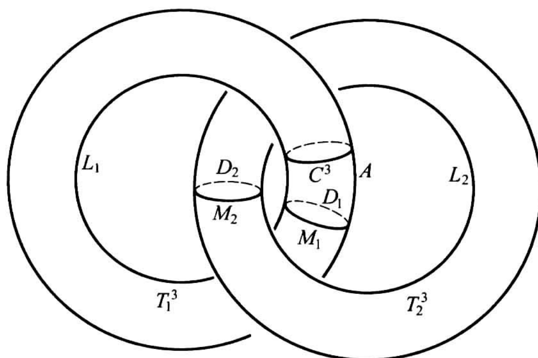


FIGURE IV.2.D

Let M_i be a meridional circle on Bd T_i^3(M_i bounds a disk cross section D_i of T_i^3 but no disk in Bd T_i^3) and L_i be a longitudinal circle on Bd T_i^3(L_i bounds a disk in R^3 - Int T_i^3 but not on Bd T_i^3) as shown in Figure IV.2.D. Let h be a homeomorphism of Bd T_1^3 onto Bd T_2^3 that takes M_1 onto L_2 and L_1 onto M_2. Then

$$S^3 = {}_tT_1^3 \cup_h T_2^3$$

The topology of $T_1^3 \cup_h T_2^3$ is as described in § IV.2.B.

This model of S^3 is a disguised version of the union of two balls given in § IV.2.B. If one thickens the disk D_1 to get a pillbox (solid cylinder) C^3 whose lateral surface is an annulus A going meridionally around Bd T_1^3, the closure of $T_1^3 - C^3$ is a 3-ball B_1^3, i.e.,

$$B_1^3 = \overline{T_1^3 - C^3}$$

Also

$$B_2^3 = C^3 \cup_{h|A} T_2^3$$

is a 3-ball and

$$S^3 = {}_tT_1^3 \cup_h T_2^3 = {}_tB_1^3 \cup_{h'} B_2^3$$

where

$$h' = \begin{cases} h \text{ on Bd } T_1^3 - A \\ \text{Identity on bases of } C^3 \end{cases}$$

Other models of S^3 are given by Bing in[$\mathbf{B_{26}}$].

IV.2.E. S^3 *is a special sort of* 3-*manifold*. An *n-manifold* is a separable metric space each point of which has a neighborhood homeomorphic to Euclidean n-space R^n. An *n-manifold-with-boundary* is a separable metric space each of whose points has a neighborhood whose closure is topologically equivalent to an n-simplex. Hence, an n-



manifold is an n-manifold-with-boundary but not conversely. Some authors call a compact n-manifold (without boundary) a *close manifold*.

Note that S^3 is compact connected 3-manifold. However, there are many other compact connected 3-manifolds topologically different from S^3. It is easy to ask what additional restrictions placed on a 3-manifold makes it a 3-sphere. The answers may be elusive.

Bing has shown [**B₁₀**] that a compact connected 3-manifold is a 3-sphere if each simple closed curve in it lies in a topological 3-ball in it. It is known [**B₄**] to be S^3 if it has a *sequentially unicoherent triangulation* T—the 3-simplexes in T can be ordered Δ_1^3, $\Delta_2^3, \cdots, \Delta_n^3$ such that for $2 \leqslant j \leqslant n$, $\Delta_j^3 \cap (U_{i<j}\Delta_i^3)$ is connected. These results are discussed in Chapter **XVIII**. The Poincaré conjecture is that a compact connected 3-manifold is S^3 if it is 1-connected as defined in **§ IV.4**, but the result is in doubt. The paper [**W₄**] contains a retraction of a mistake [**W₃**]. It was once claimed [**P₄**] that a compact connected triangulated 3-manifold was topologically S^3 if it had the property that 1-cycles in it (as discussed in Chapter **VII**) bounded in it, but that error was soon rectified [**P₅**].

IV.3. Alexander horned sphere. In the early part of this century [**A₁**] it was erroneously claimed that if S^2 is a 2-sphere in R^3, then there is a homeomorphism of R^3 onto itself that takes S^2 onto the boundary of a round ball. This would have implied that each 2-sphere in R^3 is tame and that the Schoenflies theorem generalizes to R^3. It was perhaps in trying to tighten a "proof" of this claim that Alexander discovered the Alexander horned sphere (partially shown in Figure IV.3) and attacked a PL version of the Schoenflies theorem for R^3. See [**A₃**, **A₂**].

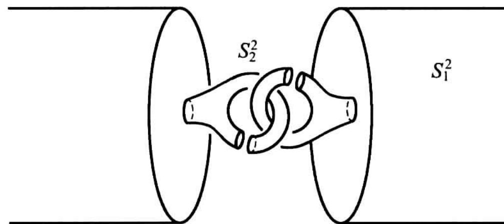


FIGURE IV.3

The Alexander horned sphere shown in Figure IV.3 can be described as the limit of a sequence of tame 2-spheres S_1^2, S_2^2, ... The first of these tame 2-spheres S_1^2 is obtained by taking a long 2-dimensional cylinder closed at both ends and bending it until the ends are near each other and parallel. Figure **IV.3** shows the ends of the cylinder

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

but not the central part. The second approximation S_2^2 is obtained by pushing tubes out of each end until they almost hook as shown. Further approximations $S_3^2, S_4^2, \ldots$, are obtained by pushing out additional tubes, pushing out additional tubes, etc. The resulting limit has the property that although it is topologically equivalent to the boundary of a round ball, the unbounded complementary domain of it is not topologically equivalent to the exterior of a round 3-ball.

EXERCISE IV. 3. A. Describe a homeomorphism between an Alexander horned sphere of Figure IV.3 and a round 2-sphere.

Solution. Let D_0, D_1 be two small mutually disjoint disks on a round 2-sphere S^2; $D_{i,j}(i=0,1; j=0,1)$ be four very small mutually disjoint disks with $D_{ij} \subset \text{Int } D_i$; D_{ijk} eight very very small disks; . . . . A homeomorphism h of the Alexander horned sphere A described above onto S^2 sends $A \cap S_i^2$ onto $S^2 - \bigcup \text{Int } D_i$, $A \cap S_2^2$ onto $S^2 - \bigcup \text{ Int } D_{ij}$, $A \cap S_3^2$ onto $S^2 - \bigcup \text{ Int } D_{ijk}, \ldots$ If $p \in A - \bigcup S_i^2$, let U_t be the component of $A - S_t^2$ containing p and D_t be the $D_{ij\ldots k}$ bounded by $h(\text{Bd } U_t)$. Then $h(p) = \bigcap D_t$.

EXERCISE IV.3.B. The set of points of the Alexander horned sphere not in any of the approximating S_i^2's is called "*the tips of the horns*". To what is this set topologically equivalent?

Answer. A Cantor set.

The union of the Alexander horned sphere and its bounded complementary domain is a crumpled cube. We could have modified our description of the Alexander horned sphere A as the limit of S_t^2's to get the crumpled cube bounded by A as the intersection of handlebodies $H_1, H_2, \ldots$

EXERCISE IV.3.C. Express $A \cup \text{Int } A$ as the intersection of handlebodies.

Answer. Let H_1 be the solid torus in Figure **IV.3.C.** It is topologically equivalent to a doughnut or the cartesian product $B^2 \times S^1$. Let C_1 be a thin meridional slice of H_1 as shown in **Figure IV.3.C.** The handlebody H_2 is obtained from H_1 by removing C_1 and adding back the ends of C_1 together with two hooked eye bolts in C_1 as shown where one of the eye bolts fits onto one end of C_1 and the other the opposite ends. An *eye bolt* is a solid torus on a stem.

To get H_3, one removes a meridional slice from each eye bolt and makes a replacement as before. The procedure is continued. Then $A \cup \text{ Int } A = \bigcap H_i$.

Crumpled cubes bounded by horned spheres come in many shapes and sizes. Figure IV.3.D shows the emblem on the front of many T-shirts. Figure IV.3.E shows a ver-



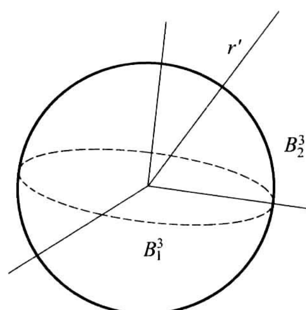
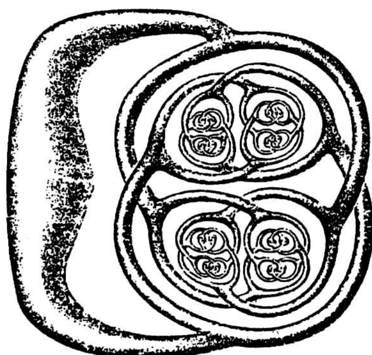


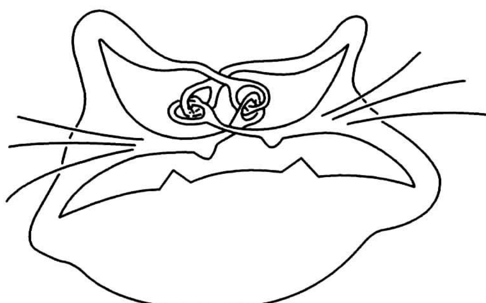
FIGURE Ⅳ.3.C

sion suggested by Bill Eaton Each of the examples described by Figures Ⅳ.3.D, Ⅳ.3.E has the property that there is a homeomorphism $h: R^3 \to R^3$ that takes Alexander's original example onto the examples. Each suggests that the Schoenflies theorem about 1-spheres in R^2 does not extend to 2-spheres in R^3.



Reprinted with permission of the publisher from Bing, R., *Some aspects of the topology of* 3-*manifolds related to the Poincare conjecture*, Lectures on Modern Mathematics vol. 2, edited by T. L. Saaty, John Wiley & Sons, Inc., New York, 1964, p. 117.

FIGURE Ⅳ.3.D



Alexander's Cat

FIGURE Ⅳ.3.E

IV.4. Simple connectivity. We examine some properties that can be used to show that Alexander's horned sphere is wild.

A set X is 1-*connected* if each map of Bd Δ^2 into X can be extended to a map of Δ^2 into X. A set which is 1-connected is also called *simply-connected*. If $f: Y \to X$, we say that f is *homotopic to a constant* if there is a map $g: Y \times [0,1] \to X$ such that $g(y \times 0) = f(y)$ and $g(Y \times 1)$ is a point. We say that a set $Y \subset X$ *can be shrunk to a point in* X if the identity map of Y into X is homotopic to a constant. We regard t as a parameter and Y as having been shrunk to $g(Y \times t)$ as time t. A necessary and sufficient condition that a set Y can be shrunk to a point in a space X is that Y is the base of a singular cone in X.

THEOREM IV.4.A. *The complement of the crumpled cube bounded by the Alexander horned sphere is not simply-connected.*

This theorem follows from **Theorem VI.10.C** from **Chapter VI.**

COROLLARY IV.4.B. *The Alexander horned sphere is wild.*

EXERCISE IV.4.C. Give a homeomorphism of a round 3-ball onto the union of the Alexander horned sphere and its bounded complementary domain.

Solution. Let X_1, X_2 be two end slices of a 3-ball B^3 as shown in **Figure IV.4.B.** A description of a suitable homeomorphism h can be started by saying that h sends the part of B^3 between X_1 and X_2 to $H_1 - C_1$ of **Figure IV.3.B.** Let X_{i1}, X_{i2} be small mutually disjoint end slices of B^3 in $X_i (i = 1, 2)$ as shown. Extend h to take $B^3 - \bigcup X_{ij}$ to H_2 minus two meridional cross sections of handles of H_2. If this procedure is continued one gets a homeomorphism h from most of B^3 to most of the crumpled cube. It is not yet defined on the Cantor set$(\bigcup X_i) \cap (\bigcup X_{ij}) \cap (X_{ijk}) \cap \cdots$ but this set is sent by h onto the tips of the horns. Had we used $C_{1ij\cdots k}$'s to designate the thin meridional slices in the H's, one could have defined the image of this Cantor set to be $C_1 \cap (\bigcup C_{1i}) \cap (\bigcup C_{1ij}) \cap \cdots$

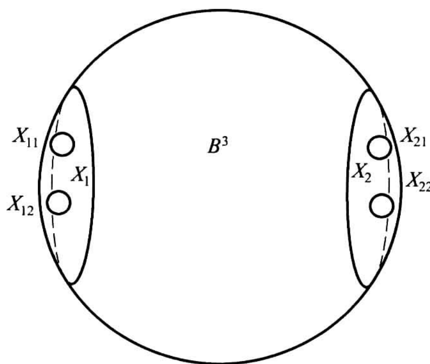


FIGURE IV.4.B

The above exercise shows it is not the interior of this crumpled cube that topologically distinguishes the cube from a round 3-ball. However, the exterior of this crumpled cube is not simply-connected. The exterior of a tame cube is simply-connected.

EXERCISE IV.4.D. If B^3 is a round 3-ball in R^3 and P is a polygon in $R^3 - B^3$, describe a map f that shows that P can be shrunk to a point in $R^3 - B^3$.

Solution. Let g be a PL map of Δ^2 into R^3 such that g takes Bd Δ^2 homeomorphically onto P. Let $p_0 \in B^3 - g(\Delta^2)$ and B^+ be a round 3-ball such that $B^3 \subset \text{Int } B^+ \subset R^3 - P$. Let ϕ be the projection from p_0 of $B^+ - p_0$ onto Bd B^+. A suitalbe map f: $\Delta^2 \to R^3 \to B^3$ can be described as follows: $f(x) = g(x)$ if $g(x) \in R^3 - B^+$; $f(x) = \phi g(x)$ if $g(x) \in B^+$.

IV.5. Solid Alexander horned sphere. If an Alexander horned sphere A is placed in S^3 rather than in R^3 and U is the nonsimply-connected complementary domain of A, any set homeomorphic to $A \cup U$ is called a *solid Alexander horned sphere*. It is a crumpled cube that is topologically different from a 3-ball.

We could have defined a solid Alexander horned sphere in R^3 rather than in S^3 if we had defined the horned sphere inside out with the tubes reaching in instead of out. We give such a description that has been useful.

Let C be a two-dimensional right circular cylinder in R^3 with bases D_1, D_2. See Figure IV.5. Two mutually exclusive disks in each $D_i (i = 1,2)$ are replaced by the surfaces of tubes T_{i1}, T_{i2}, and disks D_{i1}, D_{i2} as shown in Figure IV.5 where D_{i1}, D_{i2} are the bases of a right circular cylindr C_i and $D_1 \cup T_{11} \cup C_1 \cup T_{12}$ is hooked to $D_2 \cup T_{21} \cup C_2 \cup T_{22}$ as shown.

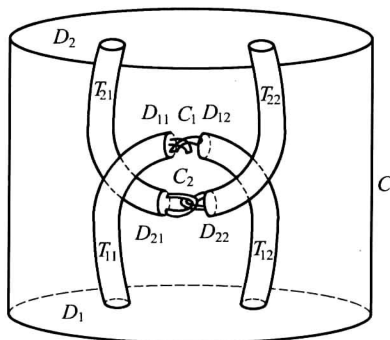


FIGURE IV.5

Disks in the bases of the cylinder $C_i (i = 1,2)$ are replaced by the surfaces of tubes $T_{i11}, T_{i12}, T_{i21}, T_{i22}$ and disks $D_{i11}, D_{i12}, D_{i22}$ as before. The process is continued to get the horned sphere M. We use M_1 to denote the part of M that is the closure of

the part of M on the exterior of $C_1 \cup C_2$. Then $M_1 = C -$ (union of four disks, two from each of D_1, D_2) $\cup$ (union of four tubes, T_{ij} ($i = 1,2; j = 1,2$)). Likewise, M_i denotes the closure of the part of M on the exterior of $\cup C_{j_1 j_2 \cdots j_i}$. It is topologically equivalent to a 2-sphere minus 2^{i+1} open disks. Let M_0 be the Cantor set $M - \cup M_i$. It is the tips of the horns of M.

Although M is homeomorphic to a round 2-sphere, its bounded complementary domain is not simply-connected. We call the complementary domain of M that is not simply-connected the bad complementary domain of M. Note that the union of M and its bad complementary domain U is a solid Alexander horned sphere.

IV.6. Peculiar involutions. We digress from our treatment of 3-manifolds to list some interesting properties of crumpled cubes that will not be considered in depth in this book. A result by Bing[**B5**] that we shall not prove but which has had applications is the following.

THEOREM IV.6.A. *The union of two solid Alexander horned spheres sewed together with an identity-like homeomorphism on their boundaries is topologically S^3 in the sense that there are two topological solid Alexander horned spheres K_1, K_2 in S^3 and a homeomorphism $h: K_1 \to K_2$ such that*

$$S^3 = K_1 \cup K_2$$

$$\text{Bd } K_1 = \text{B } K_2 = K_1 \cap K_2, \text{and}$$

$$h \mid \text{Bd } K_1 = \textit{identity}$$

It would be interesting to be able to visualize a picture of Bd K_1 = Bd $K_2 = K_1 \cap K_2$ of Theorem IV.6.A. The tips of the horns from one side would look like the tips of the horns from the other side.

Smith has shown[**S4**] that for any orientation preserving periodic homeomorphism of S^3 onto itself, the fixed point set is either null, a 1-sphere, or S^3 itself; while for any orientation reversing periodic homeomorphism, the fixed point set is either a 0-sphere or a 2-sphere. He also has shown that if h is a periodic homeomorphism of S^3 whose fixed point set is a 2-sphere, then h is an *involution* (homeomorphism of period 2). Theorem IV.6.A was the key link in showing that there is an involution of S^3 onto itself whose fixed point set is a wild 2-sphere.

THEOREM IV.6.B. *There is an involution S^3 whose fixed point set is a wild 2-sphere.*

A homeomorphism of R^n onto itself is a *rigid motion* if it is an isometry. Examples





of rigid motions of R^3 are translations, rotations about lines, reflections through planes, lines, or points, and compositions of such motions. A homeomorphism h is *equivalent to a rigid motion* if there is a homeomorphism f and a rgid motion r such that $h = f^{-1} rf$. In general, two homeomorphisms h_1, h_2 are *equivalent* if there is a homeomorphism f such that $h_1 = f^{-1} h_2 f$.

THEOREM IV.6.C. *There is an orientation reversing involution of R^3 which is inequivalent to any rigid motion.*

Once it had been shown that one could obtain S^3 by sewing two solid Alexander horned spheres together with an identity-like homeomorphism on their boundaries it was found that there were many other crumpled cubes—each topologically different from the others—such that if one took any of them and sewed it onto a topological copy of itself with an identity-like homeomorphism on their boundaries, then S^3 was obtained. See Alford[**A₇**], Bing[**B₂₀**], Montgomery and Samehon[**MS**], and Montgomery and Zippin [**MZ**]. Other sewings of crumpled cubes are given by Ball[**B₁**], Bing[**B₈**], Cannon [**C₆**], Casler[**C₇**], Daverman and Eaton[**DE₁**, **DE₂**], Eaton[**E₂**], and Martin[**M₁**].

THEOREM IV.6.D. *There is an uncountalbe family of mutually inequivalent orientation reversing involutions of R^3 onto itself.*

IV.7. Antoine's necklace. Consider in R^3 a decreasing sequence of sets T_1, $\bigcup T_{1i}$, $\bigcup T_{1ij}$, $\cdots$, such that T_1 is a topological solid torus; $\bigcup T_{1i}$ is the union of a collection of topological solid tori $\{T_{1i}\}$ such that these solid tori form a chain in Int T_1 that go around T_1 as shown in Figure IV.7.A; $\bigcup T_{1ij}$ is the union of a collection of solid tori $\{T_{1ij}\}$ such that for each T_{1i_0}, the T_{1i_0j}'s form a chain which go around T_{1i0}; $\bigcup T_{1ijk}$ is the union of topological tori T_{1ijk}; ... The limit as n increases without limit of the mesh of $\{T_{1ij\cdots n}\}$ is 0. The *mesh* of a collection of sets in a metric space (like the holes in a screen wire) is the least upper bound of the diameters of the objects. The intersection of the sets $\bigcup T_1$, $\bigcup T_{1i}$, ..., is called an Antoine's necklace.

Although Figure IV.7.A shows four T_{1i}'s in T_1, we do not make this restriction. For all we care, the number of links in the chains may vary except that in each case it exceeds two.

EXERCISE IV.7.A. Show that irrespective of how the number of links of the chains vary, there is a homeomorphism h of the resulting Antoine's necklace onto a Cantor set.

Solution. In building a Cantor set on an interval one is not required to remove an

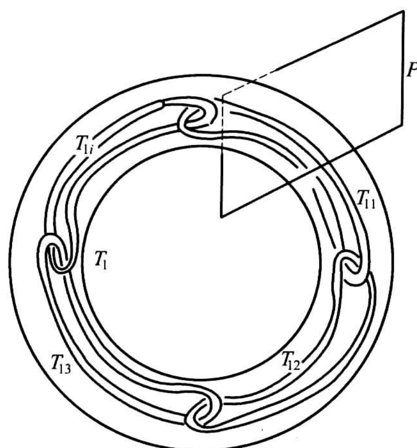


FIGURE IV.7.A

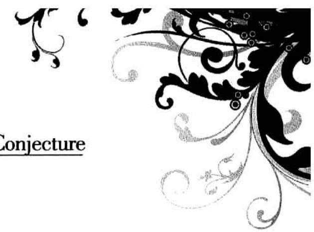
open middle third at each stage and be left with two pieces, but instead could remove several open holes and be left with several small pieces. Divide an interval into $2n_1 - 1$ equal pieces where n_1 is the number of tori in $\{T_{1i}\}$. Delete the interiors of the even pieces and agree that h sends T_{1i} into the ith remaining piece. As a next step divide the ith remaining piece into $2n_{i2} - 1$ pieces where n_{i2} is the number of tori of $\{T_{1ij}\}$ in T_{1i}; remove the interiors of the even ones and let h send T_{1ij} into the jth of the remaining. The description of h is continued. Although no step reveals precisely where a point goes, the limit does.

Antoine[**A₈**]described the necklace to illustrate the fact that even compact 0-dimensional sets in R^3 which are topologically alike might have topologically different complements. He did not actually prove that the complement of Antoine's necklace was not simply-connected but he assumed this. It was shown by Coelho [**C₉**] that the complements of certain Antoine's necklaces were not simply connected.

Someone has asked why the set is called an Antoine's necklace. Perhaps the answer is that it was first described by Antoine[**A₈**]. Like a necklace its components are small and hold together. If it were slung over one's arm, not a point would drop to the floor. Like a string of beads, it is a thing of beauty with mostly aesthetic value. It has been useful to the mathematician as an example to illustrate certain topological properties.

It may be convenient to think of the $T_{1ij\cdots n}$'s as tubular neighborhoods with small cross radius of smooth simple closed curves, but we only required that they be topologically equivalent to such objects and not that they be thin. We did require that the mesh





of the $\{T_{1ij\cdots n}\}$'s go to zero as the length of $1ij\cdots n$ increased without limit.

We examine the complement of Antoine's necklace by considering the intersections of PL surfaces. Recall from § Ⅱ.3 that two complexes(X_1, T_1), (X_2, T_2) in R^3 are in *general position* (GP) if not vertex of $T_i(i=1,2)$ lies on a vertex, 1-simplex or 2-simplex of $T_j(i \neq j)$ and no 1-simplex of T_1 intersects any 1-simplex of T_2. If(X_1, T_1), (X_2, T_2) are not GP, slight pushes at vertices of T_2 give a linear homeomorphism of $(X_2, T_2) \to (X_3, T_3)$ such that (X_1, T_1) and (X_3, T_3) are GP. If X_1, X_2 are geometric complexes rather than complexes we say X_1, X_2 are GP if X_1, X_2 have triangulation T_1, T_2 such that (X_1, T_1), (X_2, T_2) are GP. Other versions of general position are found in § Ⅸ.1.

Before showing that Antoine's necklace is wild we borrow some concepts related to Chapter VI. Figure Ⅳ.2.D shows two linked solid tori T_1^3, T_2^3. For each $i=1,2$, M_i is a meridional simple closed curve on Bd T_i^3 and L_i is a longitudinal simple closed surve on Bd T_i^3.

If f is the map of a simple closed curve J into Bd T_i^3, we say that $f(J)$ *circles* Bd T_i^3 *meridionally* (or *longitudinally*) if it crosses L_i (or M_i) algebraically a nonzero number of times. The following result follows from Theorem Ⅱ.3.C and the techniques of Chapter Ⅵ.

THEOREM Ⅳ.7.B. *Suppose T_1^3, M_1, T_2^3, L_2 of Figure* Ⅳ.2.D *are PL and f is a PL map of Δ^2 into R^3 such that $f(\mathrm{Bd}\Delta^2) \subset$* Bd T_1^3 *and $f(\Delta^2)$ is GP with respect to* Bd T_2^3. *If $f(\mathrm{BD}\ \Delta^2)$ circles* Bd T_1^3 *longitudinally, then Δ^2 contains a PL disk D^2 such that $f(\mathrm{Bd}\ D^2)$ circles T_2^3 meridionally.*

THEOREM Ⅳ.7.C. *If f is a map of a disk Δ^2 into R^3 that sends* Bd Δ^2 *homeomorphically onto P of Figure* Ⅳ.7.A, *then $f(\Delta^2)$ intersects the Antoine's necklace suggested by Figure* Ⅳ.7.A.

PROOF. Suppose contrariwise that there is a map f of Δ^2. The $f(\Delta^2)$ intersects only a finite number of the Bd $T_{1ij\cdots n}$'s. We suppose with no loss of generality that each Bd $T_{1ij\cdots n}$ is PL, f is PL, and $f(\Delta^2)$ is GP with respect to each Bd $T_{1ij\cdots n}$.

It follows from Theorem Ⅲ.3.C and techniques in Chapter VI that Δ^2 contains a subdisk D_1 such that $f(D_1)$ circles Bd T_1 meridionally. It follows from Theorem VI.10.1 that D_1 contains a disk D_2 such that $f(\mathrm{Bd}\ D_2)$ circles one of Bd T_{11}, Bd T_{12}, Bd T_{13}, Bd T_{14} either meridionally or longitudinally. However, if $f(\mathrm{Bd}\ D_2)$ circles Bd T_{1i} longitudinally it follows from Theorem IV.7.B that D_2 contains a subdisk D_3 such that $f(D_3)$ circles Bd $T_{1(i\pm1)}$ meridionally.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

Inductively we find that at each stage $1ij\cdots n$ there is a disk D in Δ^2 and $T_{1ij\cdots n}$ such that $f(D)$ circles Bd $T_{1ij\cdots n}$ meridionally. Hence, $f(D)$ intersects Antoine's necklace suggested by Figure IV.7.A.

The arguments for Theorem IV.7.C applied equally well to Antoine necklaces that had three or more $T_{1ij\cdots n}$'s at each stage in the T at the preceding stage as suggested by Figure IV.7.A. We note another property that distinguishes these Antoinie necklaces from tame Cantor sets.

A set X is *n-locally connected* (n-LC) at a point p if for each neighborhood U of p there is a neighborhood V of p such that each map of Bd Δ^{n+1} into $X \cap V$ can be extended to a map of Δ^{n+1} into $X \cap U$. A set is called *locally simply-connected* if it is 1-LC and *locally arcwise connected* if it is 0-LC.

THEOREM IV.7.D. *The complement of Antoine's necklace A suggested by Figure IV.7.A is not 1-LC at all points of Antoine's necklace.*

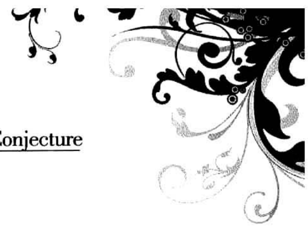
PROOF. Let D be a PL disk bounded by p of Figure IV.7.A. Assume $R^3 - A$ is 1-LC at each point of A. Let $\{V_\alpha\}$ be a collection of open sets covering $D \cap A$ such that each map of Bd Δ^2 into a $(R^3 - A) \cap V_\alpha$ can be extended to map Δ^2 into $R^3 - (A \cup P)$. Cover $A \cap D$ with a finite collection of mutually disjoint disks $D_1, D_2, \cdots, D_n$ in Int D such that each D_i lies in a V_α. Get a singular disk D' by replacing each D_i with the image of a disk into $R^3 - (A \cup P)$. Apply Dehn's lemma (see Chapter XVI) to the singular disk D' to get a disk in $R^3 - A$ bounded by P. The assumption that $R^3 - A$ is 1-LC at each point of A leads to a contradiction of Theorem IV.7.C.

EXERCISE IV.7.E. Supose C is a Cantor set on a line L in R^3 and D is a disk in R^3 whose boundary misses C. Describe a map f that shows that Bd D bounds a disk in $R^3 - C$.

Solution. Let $B_1, B_2, \cdots, B_n$ be a finite collection of mutually disjoint round 3-balls in R^3 − Bd D such that $C \subset \bigcup$ Int B_i and $p_i \in B_i - D$. Let ρ_i be a homeomorphism of B_i onto itself such that ρ_i is fixed on Bd B_i and pushes $B_i \cap D$ so close to Bd B_i that $C \cap \rho_i(B_i \cap D) = \emptyset$. Then f can be defined as follows: $f(x) = x$ if $x \in D - \bigcup B_i$; $f(x) = \rho_i(x)$ if $x \cap B_i$. We project $D \cap (B_i)$ to near Bd B_i.

IV.8. An Antoine wild sphere. Alexander [$\mathbf{A_3}$] used an Antoine's necklace to build a wild sphere. Suppose that as shown in Figure IV.8, one has an Antoine necklace on the left and a 2-sphere on the right. A disk is removed from the 2-sphere and replaced by a thin *feeler* (capped open cylinder) which runs over to Bd T_1. Then disks are removed from the part of the feeler on Bd T_1 and feelers are run inside T_1 over to each Bd T_{1i}. The process is continued. The limit is called an *Antoine wild sphere*.





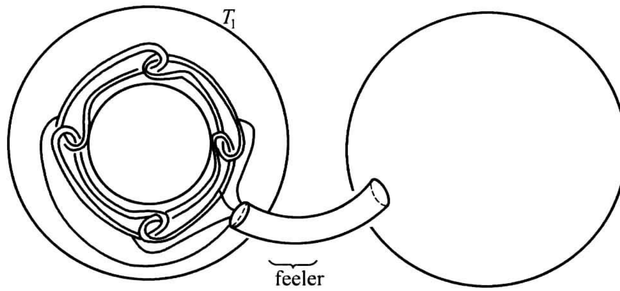


FIGURE IV.8

Perhaps Alexander's horned sphere is better known than Antoine's. Why? Perhaps it is that Antoine's is described with verbage[**A₄**] while Alexander's was given with a meaningful diagram [**A₃**]. Aspects of the following theorem are considered again in Chapter VI.

THEOREM IV.8. *The Antoine wildsphere is wild.*

IV.9. Some Fox-Artin spheres. Fox and Artin described some interesting wild spheres in[**FA**]. These spheres were the boundaries of 3-balls formed by thickening the interiors of certain wild arcs. We shall consider three of these wild spheres that have been widely used as examples.

IV.9.A. *A 2-sphere with nonsimply connected complement.* The dotted portion of Figure IV.9.A represents two solid cones with a common base $F_{-1} \cap F_1$ subdivided into a countable collection of frustrums $F_1, F_{-1}, F_2, F_{-2}, F_3, \cdots$. A more detailed frustrum F_i is shown in the right part of the figure. These frustrums are not a part of the 2-sphere but only serve as guides. The 2-sphere intersects each of these frustrums in the union of three tubular cylinders in a fashion as shown on the right of the figure. In the main part of Figure IV.9.A, we only show the center of these cylinders. If the interior of the oscillating arc pq from p to q is fattened, one obtains the 3-ball whose boundary is a wild sphere which is locally smooth except at two points.

Using techniques of knot theory such as treated in Chapter VI it can be shown that the simple closed curve J shown in frustrum F_2 of Figure IV.9.A cannot be shrunk to a point in $R^3 - pq$. Hence the 3-ball obtained by fattening pq at its interior points is wild as is its 2-sphere boundary.

IV.9.B. *A wild sphere whose complement is topologically like that of a 2-sphere.* Consider the 2-sphere shown in Figure IV.9.B. It is the union of a disk in $F_{-2} \cap F_{-1}$ and a part of the 2-sphere of § IV.9.A. Except at the left side of the sphere, we showed the center lines of the tubular cylinders rather than the cylinders themselves. We

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

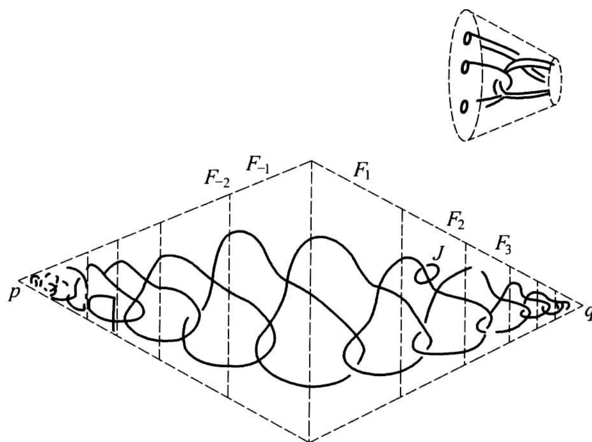


FIGURE IV.9.A

use B to denote the 3-ball bounded by the 2-sphere.

Showing that a set X is not 1-LC at a point p is a bit like playing a game with a devil. You pick a neighborhood U of p and defy the devil to pick a neighborhood V of p. Then show that no matter what V the devil picks, you can find a map f: Bd $\Delta^2 \to X \cap V$ that cannot be extended to take Δ^2 into $X \cap U$.

EXERCISE IV.9.B.a. Assume that J of Figure IV.9.A cannot be shrunk to a point in $R^3 - pq$ and that you are playing a game with a devil to show $R^3 - B$ is not 1-LC at q. How could you select U? After the devil has picked V, how could you select f: Bd $\Delta^2 \to (R^3 - B) \cap V$?

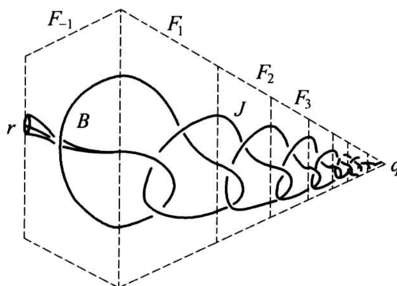


FIGURE IV.9.B

Solution. One selection (there are others) for U is the half space to the right of the plane through $F_1 \cap F_2$. After the devil has picked V pick an integer so large that $F_n \subset V$ and let P be a polygon that links B in F_n as J linked pq in F_2. If P could be shrunk to a point in $(R^3 - B) \cap U_1$, it could be shrunk to a point in $R^3 - pq$. However, it cannot be shrunk in $R^3 - pq$ because J cannot be shrunk and J can be slid (iso-



toped) to P in $R^3 - pq$.

EXERCISE IV.9.B.b. Suppose q, B are as shown in Figure IV.9.B and Bd B is locally PL except at q. Suppose $g:\Delta^2 \to R^3 - \{q\}$ is such that $B \cap g(\text{Bd } \Delta^2) = \emptyset$. Explain how to get a map $f:\Delta^2 \to R^3 - B$ which agrees with g on Bd Δ^2.

Solution. Enlarge B in $R^3 - \{q\}$ to get a 3-ball B^+ such that $B \subset \{q\} \cup \text{Int } B \subset B^+ \subset R^3 - g(\text{Bd } \Delta^2)$. Let D be a disk in Bd $B^+ - \{q\}$ that contains $(\text{Bd } B^+) \cap g(\Delta^2)$ and ϕ be a map of Δ^2 into D that agrees with g on $g^{-1}(D)$. One could use the Tietze extension theorem to get such a ϕ. Then $f(x) = g(x)$ or $\phi(x)$ according as x lies or does not lie in the same component of $\Delta^2 - g^{-1}(D)$ as Bd Δ^2.

The 3-ball B shown in Figure IV.9.B has the peculiar property that although it is wild, its complement is homeomorphic to R^3 minus the origin. A subset X of R^3 is *cellular* if it is the intersection of a decreasing sequence of 3-balls in R^3 such that each lies in the interior of the preceding.

THEOREM IV.9.B.c. *B is cellular*.

INDICATION OF PROOF. Suppose rq is a proper subarc of pq of § IV.9.A and U is an open set that contain rq. To show that B is cellular it suffices to show that there is a 3-ball in U containing rq.

To show the existence of such a 3-ball containing rq we first consider a small 3-ball B_1 in U about q so that Int B_1 contains all of rq except an arc rb and an arc ac as shown in the top part of Figure IV.9.B.c. The figure does not show all the wiggles in rb since they are not relevant.

The 3-cell B_1 is enlarged in U to a 3-ball B_2 shown in the second part of Figure IV.9.B.c so that all of rq except a short subarc $a'c$ is covered as shown. Two feelers are extended from B_1 to from B_2. Next, a tunnel is bored into B_2 so that the arc $a'cbr$ is not covered by the resulting 3-ball B_3. Then a feeler is pushed out from a' so that all of rq gets covered as shown in the lower part of Figure IV.9.B.c. It may be noted that a straight line drawn to point r from the left passes through Bd B_3 at least three times.

IV.9.C. *A wild 2-sphere whose complement is simply-connected but topologically different from the complement of a tame 2-sphere*. See Figure IV.9.C. Again we show the center lines of the tubular cylinders rather than the cylinders. Let P be the plane containing $F_{-1} \cap F_1$, X be the part of pq of § IV.9.A on or to the right of P and X' be the reflection of X through P. Then $X \cup X'$ is the union of an arc $q'aq$ and a simple closed curve K. The wild 2-sphere of this section is the boundary of the 3-ball B obtained by fattening $q'aq$ along its interior. We shall indicate why $R^3 - B$ is simply-con-

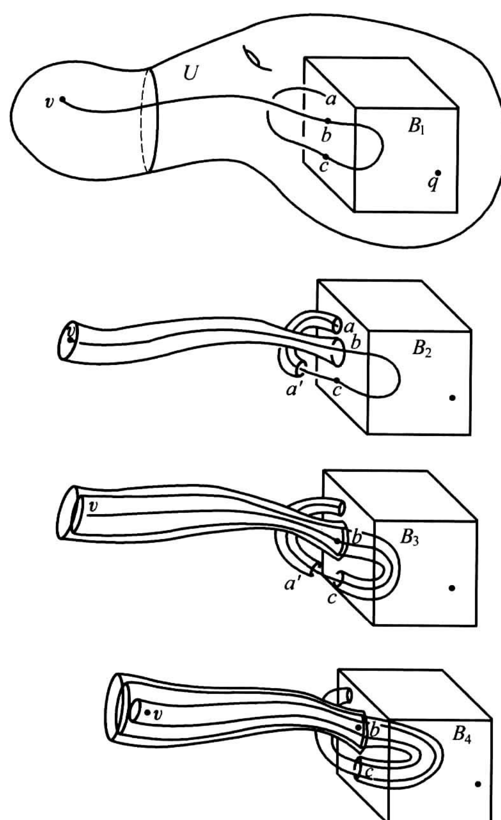


FIGURE Ⅳ.9.B.c

nected and $R^3 - B$ is topologically different from R^3 – point.

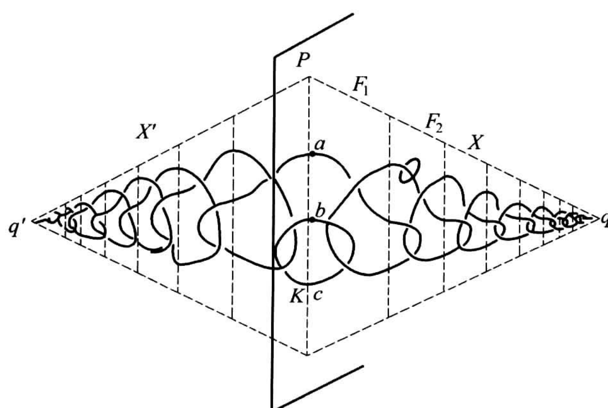
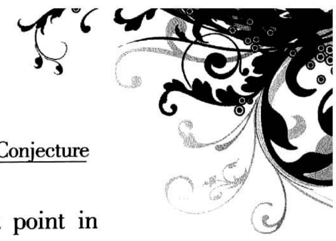


FIGURE Ⅳ.9.C





EXERCISE IV.9.C.a. Explain how to shrink J of Figure IV.9.C to a point in $R^3 - q'aq$.

Solution. Slide J along the tube about $q'aq$ until it comes into plane P near a. Then enlarge it in P so that it can be pulled over either end.

One property that makes B wild is that $R^3 - B$ is not 1-LC at q' or q. The following two theorems lead us to another propctry.

EXERCISE IV.9.C.b. Show that if J of Figure IV.9.C Can be shrunk to a point in $R^3 - (K \cup q'aq)$ then J of Figure IV.9.A can be shrunk to a point in $R^3 - pq$.

Solution. Let ϕ be the map of R^3 onto itself that is the identity to the right of the plane P of Figure IV.9.C and the reflection through P elsewhere. If $f: \Delta^2 \to R^3 - (K \cup q'aq)$ shows that J can be shrunk to a point in $R^3 - (K \cup q'aq)$ then ϕ f shows that J of Figure IV.9.A can be shrunk in $R^3 - pq$.

What we really show by this exercise is that J of Figure IV.9.C cannot be shrunk to a point in $R^3 - (K \cup q'aq)$ because we have already mentioned that J of Figure IV.9.A cannot be shrunk to a point in $R^3 - pq$.

EXERCISE IV.9.C.c. Explain why no PL ball in $R^3 - q'aq'$ contains K.

Solution. Suppose C is a 3-ball whose interior contains K and J is so small that $J \cap C = \emptyset$. It follows from Exercise IV.9.C. a that J can be shrunk to a point in $R^3 - q'aq$ with a map $f: \Delta^2 \to R^3 - q'aq$. Let ϕ be a map of Δ^2 into Bd C that agrees with f on $f^{-1}(\text{Bd } C)$. Let g be the map of Δ^2 into $\text{R}^3 - (\text{Int } C \cup q'aq)$ which is $f(x)$ or $\phi(x)$ according to whether x lies in or does not lie in some component of $\Delta^2 - f^{-1}(\text{Bd } C)$ containing Bd Δ^2. Then g shrinks J in $R^3 - (K \cup q'aq)$. This violates Exercise IV.9.C.a and the fact that J of Figure IV.9.A cannot be shrunk to a point in $R^3 - pq$.

The *lamp cord trick* is the observation that if M^3 is a hollow ball (homeomorphic image of $S^2 \times [0,1]$) and for $i = 1$ or 2, T_i is a tube joining the boundary components of BD M^3 the $\overline{\text{M}^3 - \text{T}_1} =_t \overline{\text{M}^3 - \text{T}_2}$ irrespective of whether or not T_2 is "knotted". The set $\overline{M^3 - T_i}$ resembles the closure of a room minus a light bulb on a cord as shown in Figure IV.9.C.d and the homeomorphism yielding the light cord trick comes from untying the knotted cord.

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

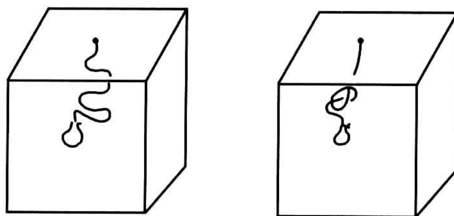


FIGURE Ⅳ.9.C.d

It follows from the lamp-cord trick that if X is a compact closed subset of R^3 and B^3 is a round 3-ball in the unbounded component of $R^3 - X$, then X lies in a tame 3-ball in $R^3 - B^3$. Figure Ⅳ.9.C.e shows C^3 as a large round ball whose interior contains $B^3 \cup X$. Since B^3 is in the unbounded component of X, there is a "lamp cord" from Bd B^3 to Bd $C^3 - X$.

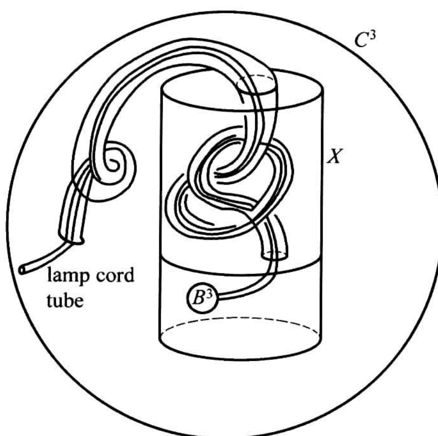


FIGURE Ⅳ.9.C.e

Ⅳ.10. Bing's hooked rug. Each of the 2-spheres we have previously considered in this chapter contained a wild arc. We now describe a 2-sphere whose wildness is so spread about that each arc in it is tame. The sphere is called *Bing's hooked rug* and is described in[B_{13}].

The hooked rug S_∞ is simultaneously the limit of a sequence of 2-spheres S_1, S_2, …, and the intersection of a decreasing sequence of 3-manifolds-with-boundary W_1, W_2, … . We describe a sequence S_1, W_1, S_2, W_2, … .

Let S_1 be a tame 2-sphere. It is our first approximation to the hooked rung S_∞. Let S_1 be the union of disks $E_1, E_2, \cdots, E_n$ such that their interiors are mutually disjoint and adjacent E_i's have an edge in common(as do E_1 and E_n). Experimentation with subdivisions of S_1 shows that we can pick E's whose diameters are small.



We thicken S_1 so that each E_i becomes a topological cube with E_i as center section. We add an eye bolt H_i to the thickened E_i so as to form a cube-with-eye-bolt T_i as shown in Figure IV.10.A. To avoid complicating the figure we do not show certain T_i in front and in back. This T_i is topologically a solid torus and the loop in each H_i loops around the stem of the next H_i (it being understood also that 1 follows n). Then $W_1 = \bigcup T_i$ is a hollow ball with handles. It may be shown that the simple closed curve J shown in Figure IV.10.A cannot be shrunk to a point in the complement of W_1.

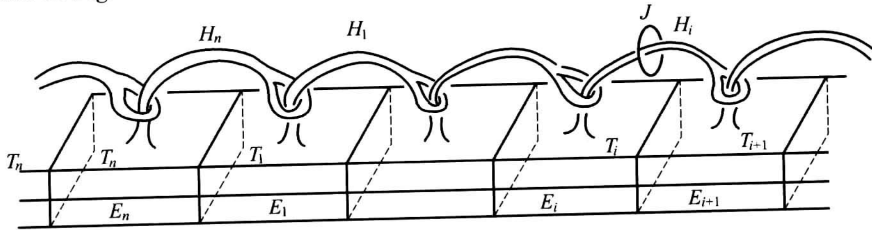


FIGURE IV.10.A

This might be an appropriate time to note hat if H_i is unhooked from the stem of H_{i+1}, J could be slipped over the right end of H_i and shrunk to a point. In fact, if instead of unhooking H_i from the stem of H_{i+1}, H_{i-1} were unhooked from the stem of H_{i+2}, J could be shrunk in a more complicated way. A part of it would need to be slipped through the hole of H_i before it finally is pulled over the right end of H_{i+1}. Indeed, if any H_j whatsoever were unhooked from the stem of its H_{j+1}, then with enough demaneuvering, J could be slipped off. While we do not need this observation in the description of this particular hooked rug, the observation is useful in building related hooked rugs.

The approximation S_2 will be obtained by replacing each disk E_i in each T_i by another disk in Bd $E_i \bigcup$ Int T_i so that this replaced disk has the same boundary as E_i. A slice is removed from the loop of each T_i to change T_i to change T_i into a topological cube K_i. The Bd E_i separates Bd K_i into two disks one of which has a hook on it. The interior of this hooked disk is pushed slightly into Int T_i so as to form a disk D_i as shown in Figure IV.10.B. The disk, except for its boundary Bd D_i = Bd E_i lies in Int T_i. The 2-sphere $S_2 = \bigcup D_i$ is the second approximation of S_∞.

Each disk D_i is given a subdivision into small disks $E_{i1}, E_{i2}, \cdots, E_{i15}$ as shown in Figure IV.10.B such that adjacent E_{ij}'s in the sequence share and edge. It is not essential that there be exactly 15 elements in the subdivision and the scheme suggested is only for convenience.

Each E_{ij} is thickened and an eye bolt added to it to from a cube-with-eye-bolt T_{ij}.

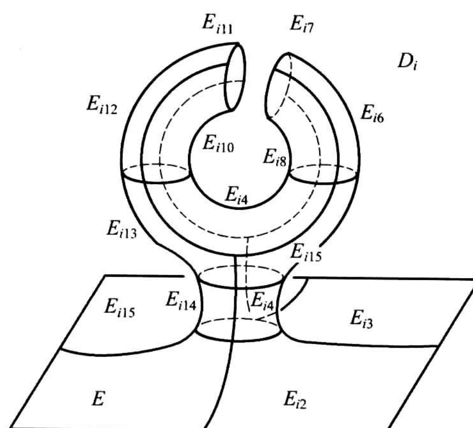


FIGURE IV.10.B

The loop in the eye bolt T_{ij} goes around the stem of the next eye bolt as shown in Figure IV.10.C. Also, the stems of T_{i7} and T_{i10} intertwine as shown in that figure so that a simple closed curve which cannot be shrunk to a point in the complement of T_i, cannot be shrunk to a point in the complement of $\bigcup T_{ij}$. The eye bolts run so close to D_i that $\bigcup T_{ij} \subset T_i$. Then $W_2 = \bigcup T_{ij}$ is a cube with handles and lies in W_1.

A disk D_{ij} is placed in each T_{ij} in the same fashion that a D_i was placed in each T_i and we define $S_3 = \bigcup D_{ij}$. Perhaps we have gone far enough for the reader to anticipate the sescription of S_∞. It is the limit of $S_1, S_2, S_3, \cdots$, and we have already described S_1, S_2, S_3. It is also the intersection of a decreasing sequence of 3-manifolds with boundaries $W_1, W_2, W_3, \cdots$, of which we have already described W_1, W_2.

To get W_{n+1} from S_{n+1} and W_n, each of the special disks D in S_{n+1} is subdivided into 15 small subdisks, each of the special subdisks is thickened into a topological cube and a topological eye bolt is added as suggested in Figure IV.10.C. The cube-with-eye-bolts T shown in Figure IV.10.C are solid even though the figure shows the eye bolts as linear and the thickened subdisks as thin. The 3-manifold-with-boundary W_{n+1} is the sum of the T's at the $(n+1)$st stage and a disk D like that shown in Figure IV.10.B is put in each sliced T. The sum of these D's is the 2-sphere S_{n+2}. It is understood that as n increases without limit, the mesh of the T's in W_n approaches 0.

For details of a proof that the limit of the S_i's $= W_1 \cap W_2 \cap \cdots$, is a wild 2-sphere such that each arc in it is tame, the reader is referred to [$\mathbf{B_{13}}$].

While we will not pursue a further study of S_∞ here, we list several of its properties.

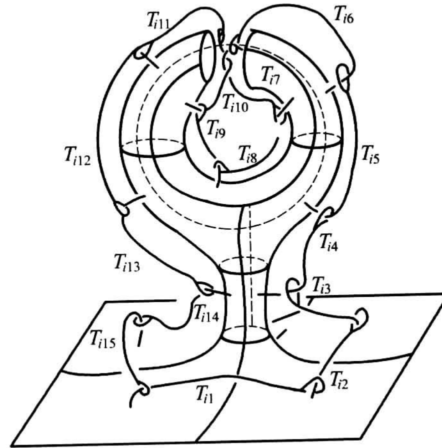


FIGURE IV.10.C

Each arc in S_∞ is tame.

$S_\infty \cup \text{Int } S_\infty$ is a 3-ball.

Ext S_∞ is not simply-connected.

Ext S_∞ is not 1-LC at any point of S_∞.

We mention two variation of S_∞ that have been used to illustrate geometric pecularities.

Instead of having H_{i15} hook the stem of H_{i1}, have it hook the stem of $H_{(i+1)1}$. The resulting 2-sphere S'_∞ has the additional property that if D is any disk in it,

$$D \cup \text{Ext } S'_\infty \text{ is simply-connected}$$

In another variation of this variation, at each stage we let the last eyebolt dangle instead of having it hook the stem of the first H at this stage. This causes Ext S'_∞ to become simply-connected. See examples of Alford [**A_6**] and Gillman[**G_2**].

CHAPTER Ⅴ THE GENERALIZED SCHOENFLIES THEOREM

V.1. The Schoenflies theorem for a 2-sphere. The Schoenflies theorem is one of the most interesting and useful theorems form plane topology. It implies that each simple closed curve in R^2 is the boundary of a disk. It is sometimes useful to consider the Schoenflies theorem on a sphere rather than in Euclidean space. If the planar version were shifted to a 2-sphere, it would say that if J is a simple closed curve on a topological 2-sphere, the the 2-sphere is the union of two disks D_1, D_2 such that $D_1 \cap D_2 = J = \text{Bd } D_1 = \text{Bd } D_2$.

Some people have the mistaken idea that since the plane was studied in high school geometry and we know so much about the plane that its study is simpler than the study of higher-dimensional topology. One of the ways to make a study easy is to restrict our attention to easy theorems. This is a way that we sometimes proceed in studying high-dimensional spaces. We are likely to attack very formidable theorems in plane topology and their proofs may be quite involved. It may be noted that the proof of the higher-dimensional Schoenflies theorem that we treat in the next § V.2 of this chapter is much easier than that given in Chapter III. However, in certain respects the theorem treated here is weaker since the hypothesis is more restrictive and the conclusions are weaker.

Early attempts were made to extend the Schoenflies theorem to higher dimensions. A claim[**A_1**] that each topological 2-sphere in R^3 bounds a topological ball was exploded by the 1924 discovery[**A_3**] of the famous Alexander horned sphere described in Chapter IV. The discovery of wild spheres showed that the Schoenflies theorem could not be extended to higher dimensions without modifications. Under what conditions can we conclude that a 2-sphere in R^3 bounds a topological 3-ball. Finding the best set of conditions is still under investigation, but we do have a satisfactory set of conditions that generalizes to all dimensions.

V.2. The generalized Schoenflies theorem. The set of all points in R^{n+1} whose distance from the origin is 1 is called the canonical n-sphere S^n. The canonical($n-1$)-sphere S^{n-1} in S^n is the set of points of S^n such that the last coordinates of each is 0. Any set homeomorphic to S^n is called a topological n-sphere. The adjective "topological" is usually dropped and we speak of a topological n-sphere as being an n-sphere.

The canonical($n-1$)-sphere S^{n-1} in S^n has a *cartesian product neighborhood* in that there is a homeomorphism h of $S^{n-1}\times[-1,1]$ into S^n such that $h(s\times 0)=s$ for each $s\in S^{n-1}$. Brown[**B_{29}**] has called S^{n-1} *bicollared* where each of $h(S^{n-1}\times(-1,0])$, $h(S^{n-1}\times[0,1))$ is a *collar*. Others say that S^{n-1} is *flat* in S^n.

Brown[**B_{29}**] showed that a topological($n-1$)-sphere K^{n-1} in R^n bounds a topological n-ball in R^n if K^{n-1} has a cartesian product neighborhood. This result is called the *generalized Schoenflies theorem*.

Brown's work was encouraged by an interesting discovery by Mazur[**M_5**] that the generalized Schoenflies theorem holds in those cases where the cartesian product neighborhood has a quality of "piecewise linearity" at one of its points. Morse[**M_{15}**] showed how to eliminate Mazur's extra requirement of "piecewise linearity". However, the following proof follows Brown's treatment[**B_{29}**].

THEOREM V.2.A (GENERALIZED SCHOENFLIFS THEOREM). *Suppose h is a*



homeomorphism of S^{n-1} into S^n such that $h(S^{n-1})$ has a cartesian product neighborhood. Then h can be extended to a homeomorphism of S^n onto itself.

We shall give a proof of this theorem for the case $n = 2$ that applies equally well in higher dimensions. We choose this setting since many readers are better acquainted with a round 2-sphere S^2 with its north pole, south pole, equator, and great circles than they are with corresponding parts of S^n.

PROOF OF GENERALIZED SCHOENFLIES THEOREM IN DIMENSION 2. We suppose S^1 is the equator of S^2, a is the north pole, and b is the south pole. We suppose that in S^2, $h(S^1 \times [-1,1])$ is a cartesian product neighborhood of $h(S^1)$ and $h(s \times 0) = h(s)$. It follows from the Jordan curve Theorem VII.6.E and invariance of domain Theorem VII.6.F of Chapter VII that $S^2 - h(S^1 \times (-1,1))$ consists of two components A, B containing $h(S^1 \times 1)$, $h(S^1 \times -1)$ respectively.

The idea for the proof is quite simple. Show that each of B, A is cellular. (A set in S^n is cellular if it is the intersection of a decreasing sequence of n-cells in S^n each of which lies in the interior of the preceding.) Then using the fact that B is celular show that there is a homeomorphism of the lower half of S^2 (part on or below equator) onto $h(S^1 \times [-1,0]) \cup B$ such that this homeomorphism agrees with h on S^1. Since A is cellular, a similar agrument shows that there is a homeomorphism of the upper half of S^2 onto $h(S^1 \times [0,1]) \cup A$.

The part of the proof that requires ingenuity is in showing that each of A, B is cellular. We need to show that for each open set U in S^2 containing B, there is a disk D_U in U such that $B \subset \text{Int } D_U$. We describe D_U as the third member of a sequence of sets $D_0, f_1(D_0), h_0(D_0) = D_U$ in S^2, the first of which is a disk D_0 such that $A \cup B \subset \text{Int } D_0$. We suppose $U \subset D_0$.

Before defining f_1 we select a map f of S^2 onto itself such that

$f(A) = a$

$f(B) = b$, and

$f(h(s \times t))$ is the point of the semicircle on S^2 from a to b through s whose last coordinate is t.

Then f restricted to $S^2 - (A \cup B)$ is a homeomorphism. We regard f as *killing* A and B. See Figure V.2.A. Let r_1 be a homeomorphism of S^2 onto itself such that $b \notin r_1 f(D_0)$ and r_1 is the identity in some neighborhood of a. Figure V.2.B illustrates such an r_1. Then

$$f_1(x) = \begin{cases} x & x \in A, \\ f^{-1}r(x), & x \in D_0 - A. \end{cases}$$

That f_1 is continuous at points of A follows from the restriction that r_1 is the identity in a neighborhood of $f(A)=a$. The map f killed A but f_1 raised it from the dead. Let us consider how $f_1(D_0)$ differs from D_0. It contains A but not B. Although $f_1(B)$ is a point of $f_1(D_0)$, B misses $f_1(D_0)$. We do not yet claim that $f_1(D_0)$ is a disk (although indeed it is). As applied to D_0, the map f_1 more nearly resembles a homeomorphism than f in that f_1 has only one nondegenerate inverse rather than two. We show $f_1(D_0)$ as shaded in Figure V.2.C. It may be shown that $f_1(U)$ is open is S^2 and f_1 is a homeomorphism on D_0-B. We have raised A from the dead. In the following step we raise B from the dead also.

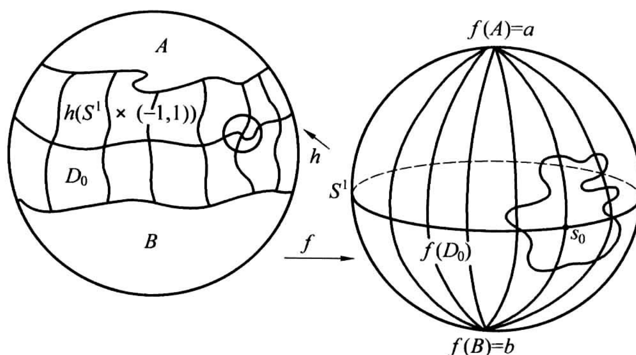


FIGURE V.2.A

We show that B is cellular by showing that for each open subset U of Int D_0 which contains B, there is a homeomorphism h_0 of D_0 into U such that $B\subset Int\ h_0(D_0)$. Then $D_U=h_0(D_0)$.

Let r_2 be a homeomorphism of S^2 onto itself such that

r_2 is the identity in some neighborhood of the point $f_1(B)$, and

$r_2f_1(D_0)\subset f_1(U)$.

Then the required homeomorphism h_0 is defined as

$$h_0(x)=\begin{cases}x, x\in B\\ f_1^{-1}r_2f_2(x), x\in D_0-B\end{cases}$$

We shall show that there is a homeomorphism of the lower half of S^2 onto $h(S^1\times[-1,0])\cup B$ that agrees with h on S^1. Since B is cellular, there is a sequence of topological disks $D_1, D_2, \cdots$, in S^2 such that

$$D_1\subset h(S^1\times[-1,0))\cup B,$$

$$\bigcap D_i=B, \text{ and}$$

$$D_{i+1}\subset \text{Int } D_i$$

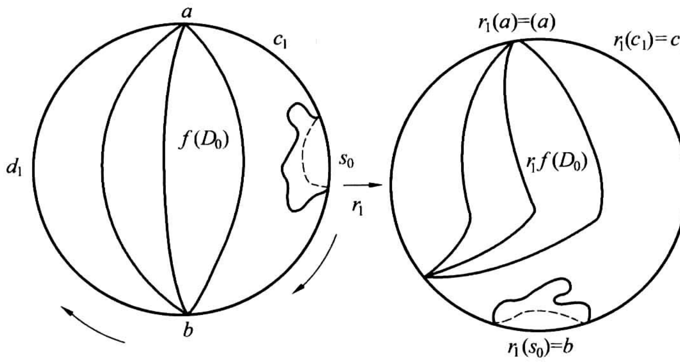


FIGURE V.2.B

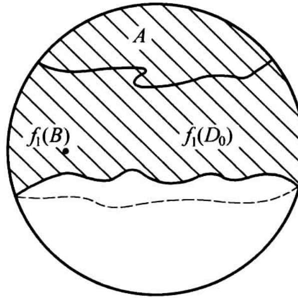


FIGURE V.2.C

Let $\phi_1, \phi_2, \cdots$, be a sequence of homeomorphisms of S^2 onto itself such that

$$\phi_1 \text{ is the identity on } S^2 - D_1$$
$$\text{diameter } \phi_1(D_2) \leqslant 1/2^1$$
$$\phi_2 = \phi_1 \quad \text{on } S^2 - D_2$$
$$\text{diameter } \phi_2(D_3) \leqslant 1/2^2$$
$$\cdots$$
$$\phi_{i+1} = \phi_1 \quad \text{on } S^2 - D_i$$
$$\text{diameter } \phi_{i+1}(D_{i+2}) \leqslant 1/2^{i+1}$$
$$\cdots$$

Let $\phi = \text{limit } \phi_i$.

A homeomorphism of the lower half of S^2 onto $h(S^1 \times [-1,1]) \cup B$ is given by ϕf^{-1}. Note that ϕf^{-1}agrees with h of S^1 since $f^{-1} = h$ on S^1 and ϕ in the identity on $h(S^1)$.

THEOREM V.2.B. *An $(n-1)$-sphere in S^n (or R^n) is flat if it has a cartesian*

product neighborhood.

The generalized Schoenflies theorem has the following application.

THEOREM V.2.C. *A compact n-manifold is a topological n-sphere if it is he union of two open n-balls.*

PROOF. Suppose the compact n-manifold M^n is the union of two open balls $\mathring{B}_1^n$, $\mathring{B}_2^n$. The compact set $M^n - \mathring{B}_1^n$ lies in $\mathring{B}_2^n$ and there is a topological n-ball B_3^n such that $M^n - \mathring{B}_1^n \subset \text{Int } B_3^n \subset B_3^n \subset \mathring{B}_2^n$ and Bd B_3^n has a cartesian product neighborhood. Since Bd B_3^n also has a cartesian product neighborhood in $\mathring{B}_1^n$, it bounds an n-ball B_4^n in $\mathring{B}_1^n$. Then M^n is a topological n-sphere since it is the union of B_3^n and B_4^n sewed together with the identity homeomorphism on their boundaries.

A subset X of Y is *collared* in Y if there is a homeomorphism h of $X \times [0,1)$ into an open subset of Y such that $h(x \times 0) = x$ for each $x \in X$. We call $h: X \times [0,1]$ a *collar* of X. Also, X is *locally collared* if for each $x \in X$, there is an open subset U of X containing x such that U is collared in Y. The generalized Schoenflies theorem has the following further application.

THEOREM V.2.D. *Suppose K^{n-1} is a topological $(n-1)$-sphere in R^n and U is a bounded componet of $R^n - K^{n-1}$. Then $K^{n-1} \cup U$ is a topological n-ball if and only if K^{n-1} is collared in $K^{n-1} \cup U$.*

PROOF. If $K^{n-1} \times [0,1)$ is the collar, it follows from Theorem V.2.A that the part of U bounded by $K^{n-1} \times 1/2$ is an *n-ball*. Then $U \cup K^{n-1}$ is this n-ball with a collar.

If f is a map of X into Y, the *set of singularities* of f is the closure in X of $\{x \in X: x \neq f^{-1}f(x)\}$. We denote this set of singularities by $S(f)$. The technique we have used to reduce singularities by raising sets from the dead may be worth formulating as a theorem.

THEOREM V.2.E. *Suppose f is a map of compact metric space X into Y, U is an open subset of $f(X)$, $A = U \cap f(S(f)) \neq 0$, and r is a homeomorphism of $f(X)$ into U such that r is the identity on an open set containing A. Let*

$$g(x) = \begin{pmatrix} x & \text{if } x \in f^{-1}(A) \\ f^{-1}rf(x) & \text{if } x \in X - f^{-1}(A) \end{pmatrix}$$

Then g is continuous and $S(g) = S(f) - f^{-1}(A)$.

ROOF. Both X and $S(f)$ are compact. If $x \in X - S(f)$ there is a neighborhood N of x such that $\overline{N} \cap S(f) = 0$. One finds successively that f, rf, and $f^{-1}rf$ are





homeomorphisms on $\overline{N}$. Also, if $x \in f^{-1}(A)$, $f^{-1}rf$ is the identity in some neighborhood of x. Finally, if $x \in S(f) - f^{-1}(A)$, rf is continuous at $rf(x) \in U - fS(f)$. Hence f^{-1} is a homeomorphism in a neighborhood of $rf(x)$.

EXERCISE V.2.F. Show that the above theorem is false if one omits the hypothesis that X is compact.

Answer. If f is a one-to-one map of a ray onto a circle and r is a rotation of the circle, $f^{-1}rf$ is not continuous at the inverse of the starting point of the ray.

V.3. The canonical collared Schoenflies theorem. Suppose S_1, S_2 are two flat $(n-1)$-spheres in S^n that are homeomorphically close(there is a homeomorphism of S_1 onto S_2 which does not move any point far). One might wonder if this homomorphism can be extended tot a homeomorphism of S^n onto itself that does not move any point of S^n far. A bit of reflection shows that if ε, δ are given(to measure"not far"and"close") before S_1, S_2, it is possible to select S_1, S_2 such that there is a δ-homeomorphism taking one onto the other but there is not ε-homeomorphism taking the ball bounded by one onto the ball bounded by the other. Figure V.3 shows two simple closed curves J_1, J_2 in R^2(or S^2) that are homeomorphically within δ of each other but the disks bounded by them are not ε close. The purpose of these remarks was to make Theorem V.3.A appear somewhat surprising. We shall be interested in conditions under which we can get a rule for extending a homeomorphism h of an$(n-1)$-sphere in S^n into S^n to a homeomorphism of S^n onto itself.

If X, Y are metric spaces with X compact, let $I(X, Y)$ denote the set of all *homeomorphism* $\{h_\alpha\}$ *of* X *into* Y metrized by

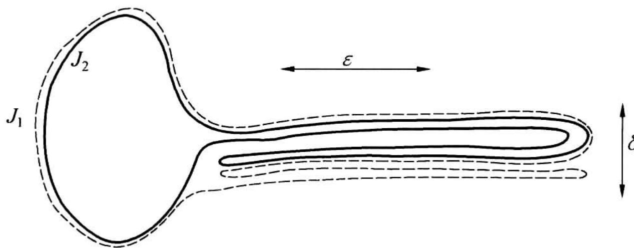


FIGURE V.3

$$d(h_\beta, h_\gamma) = \max_{x \in X} \quad d(h_\beta(x), h_\gamma(x))$$

This notation provides a vehicle for stating the following theorem.

THEOREM V.3.A(CANONICAL COLLARED SCHOENFLIES THEOREM FOR S^n). *There is a continuous map* ϕ *of* $I(S^{n-1} \times [-1,1], S^n)$ *into* $I(S^n, S^n)$ *such that each* $\phi(h_\alpha) \mid S^{n-1} \times 0 = h_\alpha \mid S^{n-1} \times 0$.

Perhaps the most spectacular application of the canonical collared Schoenflies theorem was at the third level of Kirby's remarkable diagram[**K2**] giving certain properties of maps of R^n onto itself. He used it to show that if $T^n = S^1 \times S^1 \times \cdots \times S^1$ is an n-torus and B_1, B_2 are n-cells on T^n with $B_1 \subset \text{Int } B_2$, then if h_1, h_2 are nearby elements of $I(T^n - \text{Int } B_1, T^n)$ then there are nearby elements g_1, g_2 of $I(T^n, T^n)$ such that g_i agrees with h_i on $T^n - B_2$.

We give a proof in dimension 2 that works equally well in higher dimensions. The proof consists of showing that the sets and maps used in the proof of the generalized Schoenflies theorem given inn the last section can be described in a continuous canonical fashion. The proof is complicated and the reader who is only skimming may choose to skip to § V.4 and temporarily forego the rigors of the proof.

PROOF OF CANONICAL COLLARED SCHOENFLIES THEOREM FOR S^2. We seek a canonical way for describing the maps and sets used in the proof of Theorem V.2.A. The map $h: S^1 \times [-1,1]$ is given, and A, B are precisley defined. Also, f is canonical. Instead of defining D_0 to be any disk in S^2 with $A \cup B \subset \text{Int } D_0$ we pick it to be a special one as follows. Let s_0 be the point of S^1 whose first coordinate is 1 and let D_0 be the large disk which is the complement of the small open disk with center at $h(S_0 \times 0)$ and radius $\frac{1}{3} d(h(s_0 \times 0), A \cup B)$. This makes the definition of D_0 specific and if two h's are close, their associated D_0's are close.

To pick f_1 to be canonical, we need to select an r_1 that is. Figure V.2.B shows r_1 is a variable rotation. If one wished an equation of such a rotation, one might consider the great circle on S^2 through a, s_0, and b. Let c_1, d_1 be the points of this great circle in the order a, c_1, s_0, b, d_1, a where c_1 is halfway between a and s_0 while d_1 is halfway between b and a. Then restricted to the great circle, r_1 is the identity on $d_1 a c_1$, takes $c_1 s_0$ linearly onto $c_1 b$, and takes $s_0 d_1$ linearly onto $b d_1$. Also, r_1 is defined to be the identity on the poles of S^2 with respect to this great circle. Elsewhere on S^2, r_1 is defined to take each semicircle on S^2 joining these poles linearly only another such semicircle. The second such semicircle is determined by the image of the midpoint of the first.

We proceed in this fashion to select r_1, h_0, D_U, the D_i's, and ϕ'_i canonically. This makes the homeomorphism ϕf^{-1} canonical.

Question. If B^3 is a 3-ball, it is not possible to get a map $\phi: I(\text{Bd } B^3, R^3) \to I(B^3, R^3)$ such that for each $h_\alpha \in I(\text{Bd } B^3, R^3)$, Φh_α agrees with h_α on Bd B^3 since some $h_\alpha(\text{Bd } B^3)$'s are wild from their interiors. If $I'(\text{Bd } B^3, R^3)$ is the set of elements





of $I(\mathrm{Bd}\ B^3, R^3)$ whose images are tame, is there a map $\phi: I'(\mathrm{Bd}\ B^3), R^3 \to I(B^3, R^3)$ *such that* ϕh_α and h_α agree on Bd B^3? Our proof of the canonical collared Schoenflies theorem made use of the given collar but one might think that it might be enough to know that there be a collar rather than have one selected.

Theorem V.3.A shows that it is possible to map one topological space into another subject to certain restrictions. This is a type of selection theorem. Some selection problems are quite hard. We give four examples of questions of which we know only the answers to the first and fourth. In each case we are dealing with a collection of homeomorphic compact objects in the same space and decide that the distance between two objects is the least upper bound of all ε's such that there is an ε-homeomorphism of one object onto the other.

1. If G is a collection of straight line segments in R^2, is it possible to continuously select a point from each?

2. If G is the collection of all arcs in R^2, is there a map $f: G \to R^2$ such that for each $g \in G, f(g) \in g$?

3. If G is the collection of all simple closed curves in R^2, is there such a selection map $f: G \to R^2$?

4. If G is the collection of all simple closed curves on S^2, is there such a selection map $f: G \to S^2$?

To see that the answer to Question 4 is negative, one might restrict his attention to the circles of radius $\frac{1}{4}$ on the round unit S^2. If a point is selected from a circle, the point determines a vector on S^2 from the center of the small disk bounded by the circle to the selected point. That one cannot get a continuous collection of vectors on a 2-sphere is sometimes interpreted by the layman to mean that there is a place on the earth where the wind is not blowing—you cannot comb the hair on a billiard ball.

This does not answer the question as to whether there is a map of $I'(S^{n-1}, S^n)$ into I(S^n, S^n) such that a homeomorphism in $I'(S^{n-1}, S^n)$ and its images in $I(S^n, S^n)$ agree on S^{n-1}.

V.4. Local flatness. Suppose M^{n-1} is a topological$(n-1)$-manifold in an n-manifold N^n. We say that M^{n-1} is *locally flat* if for each point $p \in M^{n-1}$ there is a neighborhood U in N^n of p and a homeomorphism of U into R^n that takes $U \cap M^{n-1}$ onto a canonial$(n-1)$-plane in R^n. Locally having a cartesian product neighborhood or being locally bicollared are essentially the same as being locally flat although these definition use the fiber approach rather than the hyperplane one. Other authors use fibers with dimensions greater than or equal to 1(see[$\mathbf{R_1}$]), but we shall ont. We learn from

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

Theorem V.4.A that S^{n-1} is tame if it is locally flat.

THEOREM V.4.A. *Any homeomorphism of one locally flat $(n-1)$-sphere in S^n (or R^n) onto another in S^n (or R^n) can be extended to a homeomorphism of S^n (or R^n) onto itself.*

The theorem follows from Theorems V.2.A and V.2.B when it is shown that an $(n-1)$-sphere has a caretsian product neighborhood if it is locally flat.

THEOREM V.4.B. *Suppose the connected $(n-1)$-manifold M^{n-1} in the n-manifold N^n is the common boundary of two components of $N^n - M^{n-1}$. Then M^{n-1} has a cartesian product neighborhood in N^n if M^{n-1} is locally flat.*

Some two-sidedness condition in the hypothesis of Theorem V.4.B is required. There is a simple closed curve J in a projective plane P^2 that does not separate P^2. Although J is locally flat in P^2, it is the center line of a Moebius band and does not have a cartesian product neighborhood.

In proving Theorem V.4.B, it is convenient to work with a component U of $N^n - M^{n-1}$ and note that M^{n-1} is locally collared in $M^{n-1} \cup U$. We shall find that the proof that M^{n-1} is collared in $M^{n-1} \cup U$ makes no use of themanifold properties of M^{n-1} so we state the results as follows.

THEOREM V.4.C. *A subset X of metric space Y is collared in Y if it is locally collared in Y.*

The following two lemmas are used in the proof of Theorem V.4.C. They show how one may adjust a collar. The adjustment was suggested by Brown in [**B$_{30}$**]. A shorter version was given by Connerly in [**C$_{10}$**].

LEMMA V.4.D. *Suppose U_1, U_2 are open subsets of X with $U_2 \subset U_1$ and h_1: $U_1 \times [0,1)$, h_2: $U_2 \times [0,1)$ are collars in Y. Then for each open set V in Y containing U_2 there is an open set V_1 in Y and a colar $h: U_1 \times [0,1)$ such that*

$$U_2 \subset V_1 \subset V$$

$$h = h_1 \quad \textit{on} \quad h_1^{-1}(Y - V), \textit{ and}$$

$$h = h_2 \quad \textit{on} \quad h_2^{-1}(V_1)$$

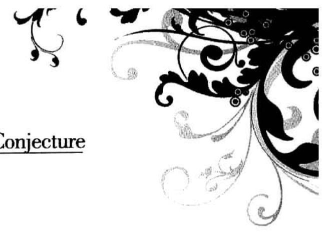
The lemma is proved by raising a blister over U_2. In describing blisters we shall consider certain maps f of U_2 into $(0,1]$ and use $h_i(U_2 \times [0,f))$ to denote the variable cartesian product $\bigcup_{x \in U_2} h_i(x \times [0,f(x)))$.

PROOF OF LEMMA V.4.D. Let f_1, f_2 be two maps of U_2 into $(0,1]$ such that

$h_1(x \times t) \in V$ and is nearer to x than to $X - U_2$ if $x \in U_2$ and $t \leqslant f_1(x)$,

$f_2 < f_1$,





$h_2(U_2 \times [0,f_2)) \subset h_1(U_2 \times [0,f_1))$.

The required open set V_1 is the blister $h_2(U_2 \times [0,f_2/2]))$. In a certain sense $V_2 = h_2(U_2 \times [0,f_2))$ is a blister twice as high. See Figure V.4.D.

Let g be the homeomorphism of Y into itself that is the identity on $Y - V_2$ and for each $x \in U_2$ sends $h_2(x \times t)$ to $h_2(x \times (t + f_2(x))/2)$ for $0 \leqslant t \leqslant f_2(x)$. In certain sense, g raises a blister over U_2 and squeezes V_2 into part of V_2 above V_1. The homeomorphism frees the blister V_1 since it sends $h_2(x \times [0,f_2(x)))$ onto $h_2(x \times [f_2(x)/2, f_2(x)))$. If $A = x \times [0,f_2(x))$, Figure V.4.D shows $h_2(A), gh_2(A), h_1(A)$, and $gh_1(A)$. Points are raised.

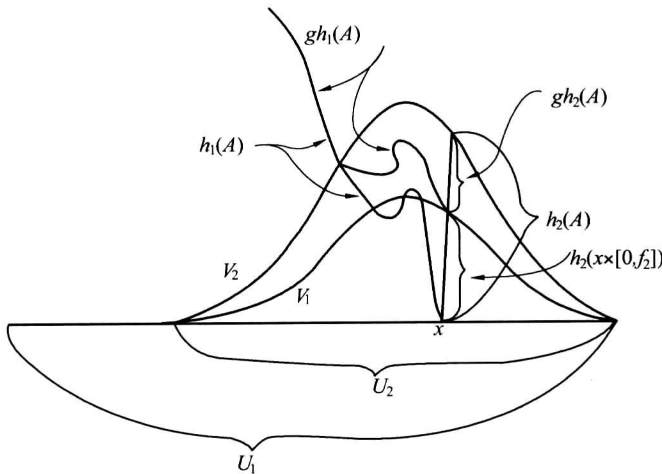


FIGURE V.4.D

The required homeomorphism $h: U_1 \times [0,1]$ *sends* $x \times [0,1]$ onto

$$h_2(x \times [0,f_2/2)) \cup gh_1(x \times [0,1))$$

where the parametrization is adjusted along $h(x \times [0,1))$ in a natural way as follows. If r_x is the linear map that sends the segment $[f_2(x)/2, f_2(x)]$ onto $[0,f_2(x)]$

$$h(x \times t) = h_1(x \times t) \quad \text{for } 0 \leqslant t \leqslant f_2(x)/2,$$

$$h(x \times t) = gh_2(x \times r_x(t)) \quad \text{for } f_2(x)/2 \leqslant t \leqslant f_2(x),$$

$$h(x \times t) = gh_2(x \times t) \quad \text{for } f_2(x) \leqslant t < 1.$$

LEMMA V.4.E. *Suppose* U_1, U_2 *are open subsets of* X *and* $h_1: U_1 \times [0,1), h_2: U_2 \times [0,1)$ *are collars in* Y *such that each* $h_1(x \times t)$ *is nearer to* x *than to* $X - U_1$. *Then there is a collar* $h: (U_1 \cup U_2) \times [0,1)$ *such that* $h = h_1$ *on* $(U_1 - U_2) \times [0,1)$ *and each* $h(x \times t)$ *is nearer to* x *than to* $X - (U_1 \cup U_2)$.

PROOF. Since Lemma V.4.D syas that we can adjust h_2 to make it agree with h_1

in a neighborhood of $(U_1 \cap U_2)\times 0$ in $(U_1 \cap U_2)\times[0,1)$, we suppose with no loss of generality that there is a map $f_1:(U_1 \cap U_2)\to(0,1)$ such that h_2 agrees with h_1 on $(U_1 \cap U_2)\times[0,f_1)$.

Let U'_1, U'_2 be disjoint open subsets of $U_1 \cup U_2$ containing $U_1 - U_2, U_2 - U_1$ respectively. Since each $h_1(x\times t)$ is closer to x than to $X - U_1$, there is a map $f_2: U_2\to(0,1)$ such that $h_1(U'_1\times[0,1))$ misses $h_2(U'_2\times[0,f_2))$ and for $t < f_2(x)$, $h_2(x\times t)$ is nearer to x than to $X - U_2$. Let f be a map of $U_1 \cup U_2$ into $(0,1)$ such that

$$f = 1 \quad \text{on } U_1 - U_2$$
$$f < f_1 \quad \text{on } U_1 \cap U_2 - (U'_1 \cup U'_2)$$
$$f < f_2 \quad \text{on } U'_2$$

Letting r_x be the linear map that takes $[0,1]$ to $[0,f(x)]$, we define h as follows

$$h(x\times t) = h_1(x\times r_x(t)) \quad \text{for } x\in U_1 - U'_2$$
$$h(x\times t) = h_2(x\times r_x(t)) \quad \text{for } x\in U_2 - U'_1$$

It is to be noted that $h_1 = h_2$ on $x\times r_x(t)$ for $x\in(U_1 \cap U_2)-(U'_1 \cup U'_2)$ since $f < f_1$.

PROOF OF THEOREM V.4.C. Let $\{U_\alpha\}$ be an open covering of X such that each U_α has a collar in Y. Since a metric space is paracompact, we can suppose with no loss of generality that $\{U_\alpha\}$ is locally finite—that is, each point of X lies in an open subset U of X such that U intersects at most a finite number of elements of $\{U_\alpha\}$. For each U_α let $h_\alpha(U_\alpha\times[0,1))$ be a collar for U_α such that each $h_\alpha(x\times t)$ is closer to x than to $X - U_\alpha$.

Theorem V.4.C follows directly from Lemma V.4.E if $\{U_\alpha\}$ is finite or countable but requires a bit of fancy footwork if $\{U_\alpha\}$ is uncountable.

Suppose $\{U_\alpha\}$ is uncountable and $W = (1,2,\ldots,\alpha,\ldots)$ is a well ordering with no last element of the index set of $\{U_\alpha\}$. For each $\beta\in W$ we let $U_\beta^+ = \cup_{\alpha<\beta}U_\alpha$. As indicated in the following paragraph, it follows from Lemma V.4.E that there is an uncountable sequence of collars $h_2: U_2^+\times[0,1), h_3: U_3^+\times[0,1),\ldots, h_\alpha: U_\alpha^+\times[0,1),\ldots$, such that each $h_\alpha(x\times t)$ is nearer to x than to $X - U_\alpha^+$ and if $\beta_1 < \beta_2$ and $x\in U_{\beta_1}^+$ but $x\notin \cup_{\beta_1=\alpha<\beta_2}U_\alpha$, then $h_{\beta_1}(x\times t) = h_{\beta_2}(x\times t)$. To define h on $x\times[0,1)$ we let U_γ be the last element of $\{U_\alpha\}$ to contain x and let $h = h_{\gamma+1}$ on $x\times[0,1)$.

We shall abbreviate $h_\alpha: U_\alpha^+\times[0,1]$ by h_α. To prove that there is an uncountable sequence $h_2, h_3,\ldots, h_\alpha,\ldots$, as mentioned in the preceding paragraph, one might proceed as follows. For each U_α^+ for which there is a collar h_α such that each $h_\alpha(x\times t)$ is closer to x than to $X - U_\alpha^+$, let W_α be a well ordering of all such collars. For two



sequences $h_2^i, h_3^i, \cdots, h_\alpha^i, \ldots (i = 1,2)$ we use the lexicographical ordering to see which sequence comes first—find the first place where the sequences differ(say in the γth slot) and then decide which of the two sequences comes first on the basis of whether or not h_γ^1 or h_γ^2 comes first in W_γ: if the two sequences agree on each place where each is defined, the shorter sequence comes first. Instead of showing that there exists a sequence $h_2, h_3, \ldots, h_\alpha, \ldots$, one has the easier task of showing that there is a unique such sequence with the additional property that $h_{\alpha+1}$ is the first sequence in $W_{\alpha+1}$ that agrees with h_α on $U_\alpha^+ - U_\alpha$. If α is a limit ordinal, h_α is automatically determined by the nature of its predecessors.

V.5. Weakness of the generalized Schoenflies theorem. A homeomorphism of R^n onto itself is *stable* if it is the composition of a finite number of homeomorphisms each of which is fixed on an open set. The Schoenflies homeomorphisms given by Theorems Ⅲ.1.A and Ⅲ.6.C are stable since they have compact support. A translation of Theorem V.2.A to R^n would say that if B_1^n, B_2^n are n-balls with locally flat boundaries in R^n, then there is a homeomorphism $h: R^n \to RR^n$ *such that* $h(B_1^n) = B_2^n$. *However there is no assurance that* h *does not shake* R^n at its very roots.

In Chapter XIV we do show that in the case where $n = 3$ and the Bd B_i^n are PL then there is a PL homeomorphism $h: R^3 \to R^3$ such that $h(B_1^n) = h(B_2^n)$ and h has compact support. The side approximation theorem for R^3 is given in Chapter XIII. As pointed out there, this result(or even the approximation theorem), can be used to show that if M^2 is a locally flat 2-manifold in R^3 and $\varepsilon > 0$, there is a homeomorphism h: $R^3 \to R^3$ such that $d(h, \text{Identity}) < \varepsilon$, h is identity outside ε – neighborhood of M^2, and $h(M^2)$ is PL. Hence if B_1^3, B_2^3 are two 3-balls with locally flat boundaries in R^3 there is a homeomorphism $h: R^3 \to R^3$ with compact support such that $h_1(B_1^3) = h_1(B_2^3)$. However, it is to be noted that the method of proof is more difficult than that used in Theorem V.2.A.

The *annulus conjecture* is that if B_1^n, B_2^n are locally flat n-balls in $\mathbf{R}^n$ with $B_1^n \subset$ Int B_2^n, B_2^n – Int B_1^n is homeomorphic to $S^{n-1} \times [-1,1]$. Theorem V.2.A does not provide a proof of the annulus conjecture since the homeomorphism h given, might not be fixed on any open set. However the techniques of Chapter III and the applications cited in Chapter XVIII can be used to establish the annulus conjecture in R^2 and R^3.

The techniques used to prove Theorem V.2.A about a locally flat$(n-1)$-sphere in R^n does not seem to extend to apply to locally flat$(n-1)$-manifolds. They apply only to$(n-1)$-spheres. However the approximation theorems given in Chapter XIII do apply to 2-manifolds in R^3.

CHAPTER Ⅵ THE FUNDAMENTAL GROUP

In Chapter VI we had need to show that certain sets are not simply-connected. In the present chapter we introduce machinery that is frequently used for these purposes. We give an algorithm for computing the fundamental group of the complement of a finite graph in R^3. The study of fundamental groups is a part of homotopy theory. However, it is not our aim to consider the broad and useful area of homotopy theory but rather to concentrate on a part of it that has had such a strong impact on the study of 3-manifolds.

Some authors regard R^3 as euclidean and denote it by E^3. To familiarize the reader with his notation, we shall occasionally use it in this chapter.

VI.1. Paths and loops. A *path* in a topological space X is a map f of the segment $[0,1]$ into X. The images $f(0)$ and $f(1)$ are called the *beginning* and *end points* of the path. They are also called the *initial* and *terminal points*.

Multiplication of paths. If f_1, f_2 are paths in X such that the end of f_1 is the beginning of f_2, the *product of* f_1 *and* f_2 (written $f_1 * f_2$) is the path

$$f_1 * f_2 = \begin{cases} f_1(2t), & 0 \leqslant t \leqslant \frac{1}{2} \\ f_2(2t-1), & \frac{1}{2} \leqslant t \leqslant 1 \end{cases}$$

In moving along the path $f_1 * f_2$ as suggested in Figure VI.1.A, one moves along the path f_1 at double speed until the end is reached and then moves along f_2 at double speed. It is to be noted that $f_1 * f_2$ is not defined unless the end of f_1 is the beginning of f_2.

Inverse. If f is a path, we call

$$f^{-1}(t) = f(1-t), 0 \leqslant t \leqslant 1$$

the *inverse* of f. As indicated in Figure VI.1.B, we are moving along f in the opposite direction.

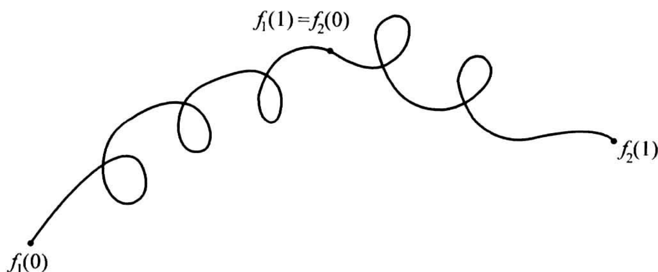


FIGURE VI.1.A

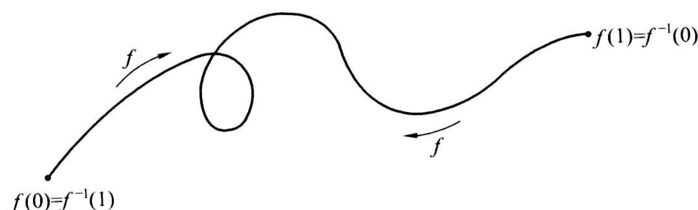


FIGURE VI.1.B

Homotopy of paths. If f_1, f_2 are two paths with the same beginnings and ends we say that f_1 *is homotopic to* f_2 *with ends fixed* (written $f_1 \sim f_2$) if there is a continuous family of maps $F_t:[0,1]\to X\,|\,(0\leqslant t\leqslant 1)$ such that $F_0=f_1$, $F_1=f_2$, $F_t(0)=f_1(0)=f_2(0)$, $F_t(1)=f_1(1)=f_2(1)$. The map is sometimes called a *homotopy*. It may be convenient to regard the homotopy as a continuous shifting of the path f_1 until it becomes the path f_2 as suggested in Figure VI.1.C. We may regard F as mapping the unit planar square with the opposite vertices $(0,1)$, $(1,1)$ into X such that $F(x,y)=F_y(x)$. Note that F is constant on each vertical side of the square. On the bottom of the square F is determined by f_1, and on the top it is determined by f_2.

Equivalence classes of paths. The above homotopy breaks the paths in X into equivalence classes. We use $[f]$ to denote the set of all paths g in X such that $f\sim g$. To see that $f\sim g$ implies $g\sim f$ one need only reflect the unit square across a horizontal segment through its center. To visualize the face that $f_1\sim f_2$ and $f_2\sim f_3$ implies $f_1\sim f_3$ one might consider the maps F and G on the unit squares showing that $f_1\sim f_2$ and $f_2\sim f_3$ respectively, squash each of these squares to one half its normal height, and then set the second squashed square on top of the first. The homotopy H showing that $f_1\sim f_3$ is given by

$$
\begin{aligned}
H_t &= F_{2t}, 0\leqslant t\leqslant \frac{1}{2}\\
&= G_{2t-1}, \frac{1}{2}\leqslant t\leqslant 1
\end{aligned}
$$

Finally, $f\sim f$.

Associativity. Suppose f_1, f_2, f_3 are paths such that the end of f_1 is the beginning of f_{i+1}, $i=1,2$. Then $(f_1 * f_2) * f_3$ and $f_1 * (f_2 * f_3)$ are defined, although they may be different as paths. Under $(f_1 * f_2) * f_3$ one considers moving along f_1 at four times normal speed, along f_2 at four times normal speed, and along f_3 at double speed. The rates of speed on f_1 and f_3 are interchanged for $f_1 * (f_2 * f_3)$. If F is defined so as to be constant on each of the slanting segments shown in Figure VI.1.D, there results a homotopy showing that $f_1 * (f_2 * f_3)\sim (f_1 * f_2) * f_3$.

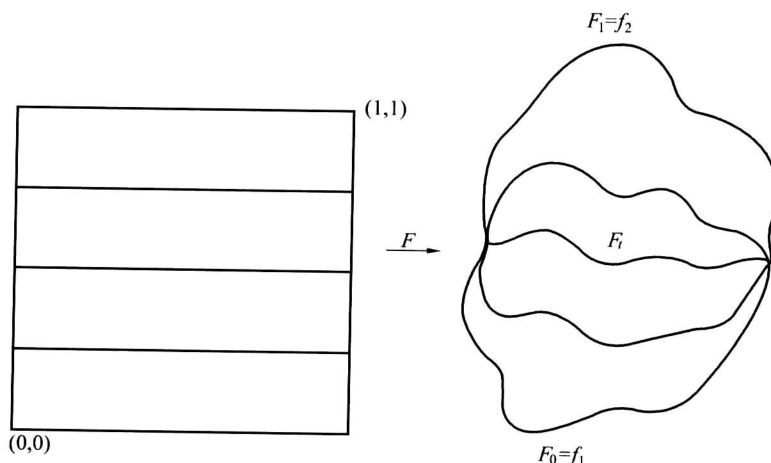


FIGURE VI.1.C

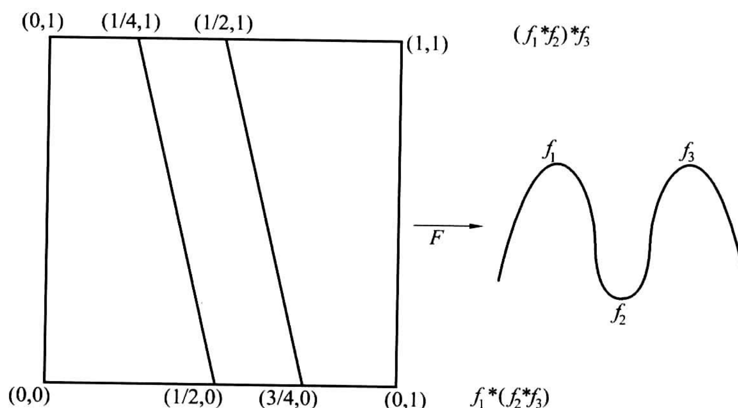


FIGURE VI.1.D

Loops. Select a point $x_0 \in X$ and designate it as a *base point*. A *loop* in X based at x_0 is a path in X that has this base point as both a beginning and an end point. See Figure VI.1.E. Note that any two loops based at x_0 can be multiplied. This puts us in a position to use equivalence classes of loops based at x_0 as elements of a group.

VI.2. The fundamental group. Let x_0 be a base point in a topological space X. We use $\pi_1(X, x_0)$ to denote the *fundamental group* of X based at x_0. An element of this fundamental group is an equivalence class of loops in X based at x_0. If $[f] \in \pi_1(X, x_0)$, $[f]^{-1}$ is the equivalence class of loops containing f^{-1}. To multiply two elements $[f]$, $[g]$ of $\pi_1(X, x_0)$ we select any two elements $f \in [f]$, $g \in [g]$, and let $[f] * [g] = [f * g]$. It can be shown that this multiplication is independent of the elements chosen in the class. It turns out that the identity element of $\pi_1(X, x_0)$ is the e-

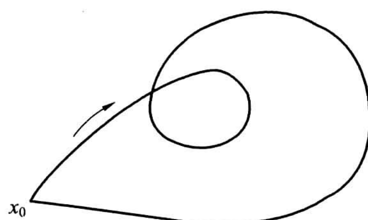


FIGURE VI.1.E

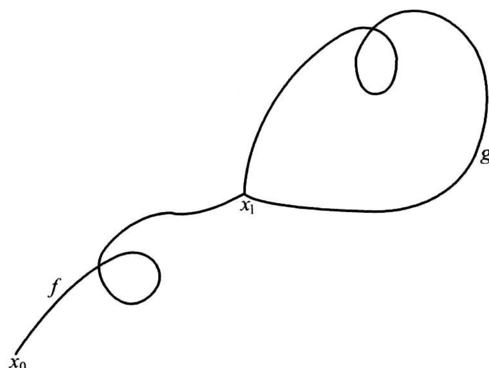


FIGURE VI.2

quivalence class containing the constant map $f[0,1] = x_0$. Multiplication is associative and π_1 is a group.

It may be shown that if there is a path f in X from x_0 to x_1, then $\pi_1(X, x_1)$ is isomorphic to $\pi_1(X, X_0)$. To get such an isomorphism one could let the loop g based at x_1 correspond to the loop $f * g * f^{-1}$ based at x_0, as shown in Figure VI.2. This correspondence between loops is extended to a correspondence between equivalence classes of loops.

When dealing with a pathwise connected space X, it is customary to ignore the base point and use $\pi_1(X)$ to designate $\pi_1(X, x_0)$. We shall frequently do this in the followinig sections.

VI.3. Graphs. A PL *finite graph* in E^3 is the sum of a finite number of straight segments such that if two of these segments intersect, the intersection is an end point of each. It is a 1-complex.

A plane P in E^3 is in *regular position* with respect to a PL finite graph G if only a finite number of lines normal to P intersect G in more than one point and each of these exceptional normal lines intersects G in precisely two points which are interior points of the segments used to describe G. We wish to alter situations as shown in the left side of

Figure VI.3.A so that they become as shown in the right side.

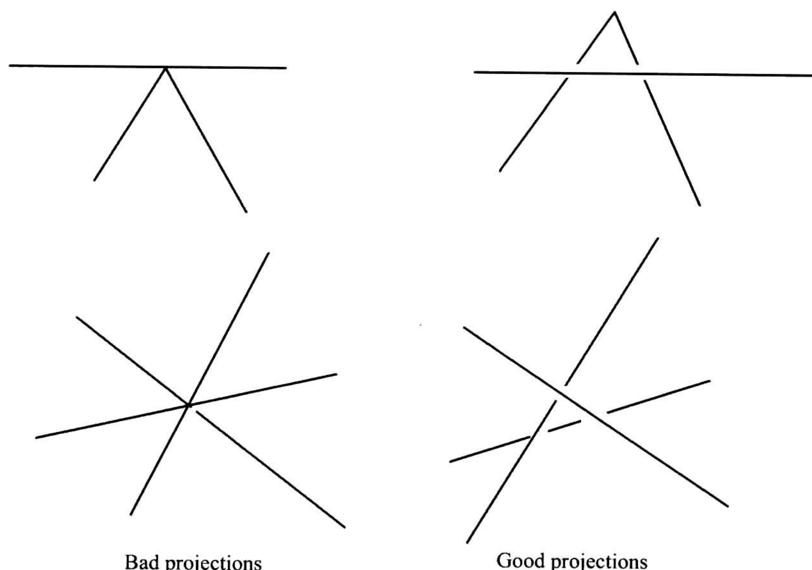


FIGURE VI.3.A

If a plane P is in regular position with respect to G, any plane parallel to P is in regular position with respect to G. If one considers the unit vectors from the origin which are normal to planes which are in regular position with respect to G, these vectors intersect the unit sphere with center at the origin in a dense subset of the unit sphere. Hence, there are many vectors such that any plane normal to this vector are in regular position with respect to G.

Consider a projection G' of G into the plane P wich is in regular position with respect to G. For convenience we regard P as horizontal so that each point of G is on a verticalline with its corresponding point in G'. If one segment of G contains a point p which is beneath a second segment we say that p is an *undercrossing point* and the first segment *passes beneath* the second. In drawing G', it is convenient to show a gap in the lower segment if one passes beneath the other. In the graph G (or its projection G') shown in Figure VI.3.B, there are five undercrossing points, three branch points and one end point.

Although for technicalities in giving proofs, one may give lip service to the use of polyhedral graphs, in practical situations one usually draws smooth graphs. The graph shown in Figure VI.3.B is shown as a smooth graph in Figure VI.3.C. In working with smooth graphs one insists that if one smooth segment of G passes underneath another,



there is actual crossing in G' and not mere tangency.

There is a theory for studying the fundamental groups of the complements of wild graphs but that is not treated in this book.

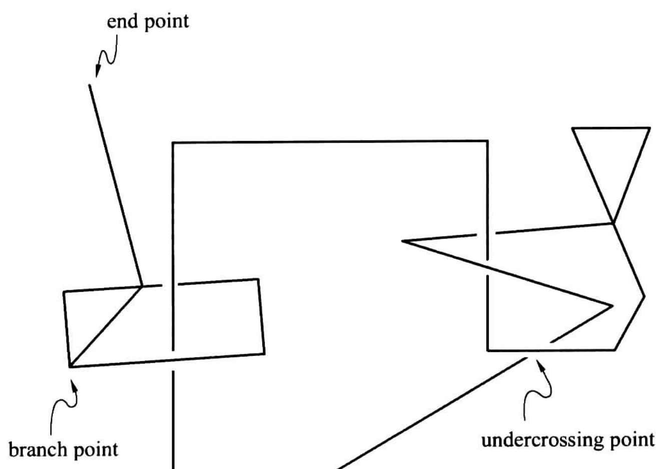


FIGURE VI.3.B

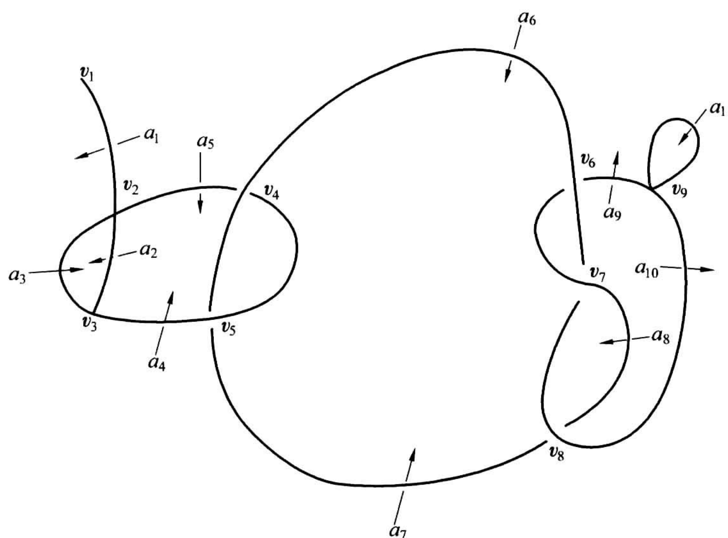


FIGURE VI.3.C

VI.4. Associative words with loops. Suppose G is a graph (either polygonal or smooth) in E^3 with a nice projection as shown in either Figure VI.3.B or VI.3.C. Let $v_1, v_2, \cdots, v_m$ be the points of G that are end points, branch points, or undercrossing points. Let $A_1, A_2, \cdots, A_n$ be the collection of open arcs and simple closed curves of G

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

such that no point of any A_i is a v_j but each boundary point of A_i is a v_j. The closure of each A_i is either an arc in G or a simple closed curve in G which intersects at most one v_j. To each A_i we associate an arrow going under A_i and designate it with a letter a_i as shown in Figure VI.3.C. It does not matter in which direction the arrows (short oriented vectors) point, but it is customary to think of moving along the A_i's of G in some direction and for the arrows to go under from the left side of A_i to the right. In most drawings, there is no need to label the v_i's and the A_i's since it is only the a_i's that are used.

Suppose x_0 is a point of $E^3 - G$ selected as a base point and f is a loop based at x_0 taking $[0,1]$ into $E^3 - G$. Figure VI.4 shows the image of such an f. We suppose that f is *normal* with respect to G in the sense that $[0,1]$ can be broken into a finite number of segments $[b_0, b_1], [b_1, b_2], \cdots, [b_{r-1}, b_r]$ $(0 = b_0 < b_1 < \cdots < b_r = 1)$ such that f is a homeomorphism on each $[b_{i-1}, b_i]$ and either no point of $f[b_{i-1}, b_i]$ is beneath G or exactly one point of $f[b_{i-1}, b_i]$ is beneath G and f crosses under a segment of G at this point but not under an undercrossing, branch, or end point of G. Although not each loop of $[f]$ has such nice undercrossing properties, for each $g \in [f]$ and each $\varepsilon > 0$ there is a piecewise linear loop f such that $d[f, g] < \varepsilon$. Pushes can change f to a piecewise linear loop in $[f]$ that is normal with respect to G.

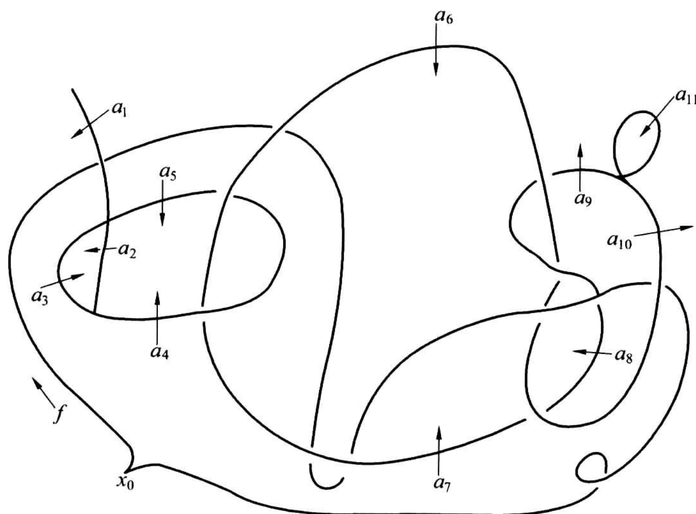


FIGURE VI.4

To each of the r pieces of $[0,1]$ we associate either a blank or a letter. If $f[b_{i-1}, b_i]$ does not cross under G, we associate a blank. If $f[b_{i-1}, b_i]$ crosses under A_j in

the direction of the arrow a_j, we associate the letter a_j; if it crosses under A_j in the opposite direction, we associate a_j^{-1}.

To the loop f we associate a word made up of letters

$$a_1, a_1^{-1}, a_2, a_2^{-1}, \cdots, a_n, a_n^{-1}$$

written as follows: We ignore blanks and first write down the letter associated with $f[b_0, b_1]$, then the letter associated with $f[b_1, b_2], \ldots$, and finally the letter associated with $f[b_{r-1}, b_r]$. In case f crosses under G we obtain a word with a finite number of letters. The word associated with the loop shown in Figure VI.4 is $a_1^{-1}a_6a_7^{-1}a_7a_{10}$. It does not contain the letters a_{10}^{-1} and a_8^{-1} since f went over G at these crossings rather than under. (Some authors borrow the convention from functions that if fg is a composite function then g is performed first. They let the last letter correspond to the first undercrossing. We find it convenint not to write backwards and use the nth letter to denote the nth undercrossing)

Let $W(f)$ be the word associated with f.

Suppose f_1, f_2 are two loops associated with the same word $W(f_1) = W(f_2) = x_1x_2\cdots x_s$. It may be shown in this case that $f_1 \sim f_2$ so f_1, f_2 belong to the same element of $\pi_1(E^3 - G)$. This is shown by the following exercise.

EXERCISE VI.4.A. Show that if $W(f_1) = W(f_2)$ then

$$f_1 \sim f_2$$

Solution. Let $y_1, y_2, \cdots, y_s$ and $y'_1, y'_2, \cdots, y'_s$ be values of $[0,1]$ such that the $f_1(y_i)$'s and the $f_2(y_i)$'s are beneath G. Since there is a homeomorphism h of $[0,1]$ onto itself so that $h(y'_i) = y_i$, we suppose without loss of generality that $y'_i = y_i$.

We start modifying f_1 by getting an f_3 such that $W(f_1) = W(f_2) = W(f_3)$ and $f_2(y_i) = f_3(y_i)$. To get f_3, we modify f_1 near the y_i's. As shown in Figure VI.4.A.a. A short arc p_iq_i of $f_1[0,1]$ about $f_1(y_i)$ is removed, and it is replaced by a folded arc that travels from p_i up vertically to near G, then travels parallel to G(but not crossing over or under G) until it is almost over $f_2(y_i)$, then descending to $f_2[0,1]$ and crosses under G with an arc of $f_2[0,1]$ containing $f_2(y_i)$, then rises vertically to near G, then runs parallel to G back to a point above q_i and then finally drops down to q_i. Since $f_3[y_{i-1}, y_i]$ has the same ends as $f_2[y_{i-1}, y_i]$ and each runs above G except at its end points, there is a homotopy of f_3 to f_2 in $E^3 - G$ that leaves each of $0, y_1, \cdots, y_s$, 1 fixed. Since $f_1 \sim f_3 \sim f_2$, $[f_1]$ and $[f_2]$ are the same element of $\pi_1(E^3 - G)$.

Although we have shown that $f_1 \sim f_2$ if $W(f_1) = W(f_2)$, one should not conclude the converse that $f_1 \sim f_2$ implies $W(f_1) = W(f_2)$. Such is not the case. It may be that

f_1 ducks under A_i as shown in Figure VI.4.A.b and picks up the letters $a_i a_i^{-1}$. In fact two loops f_1, f_2 are homotopic if $W(f_2)$ can be obtained by either cancelling or inserting the adjacent letters $a_i a_i^{-1}$ or $a_i^{-1} a_i$ any place in $W(f_1)$. In the next section we point out other conditions under which different words are associated with homotopic loops.

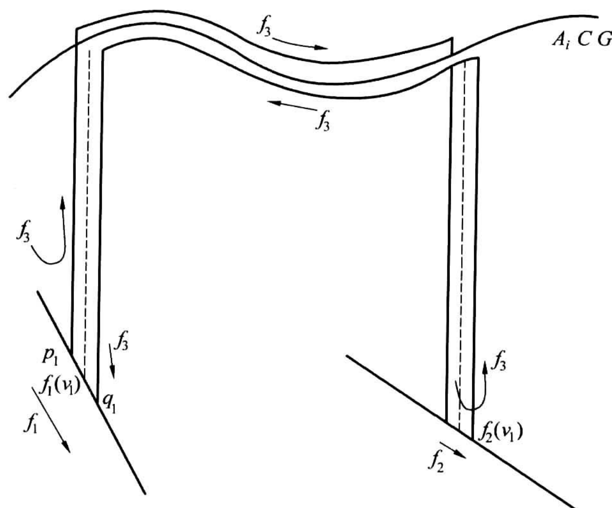


FIGURE VI.4.A.a

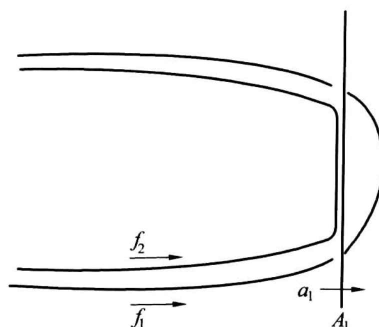


FIGURE VI.4.A.b

VI.5. Relations. We shall consider operations by which $W(f_1)$ can be changed to $W(f_2)$ if $f_1 \sim f_2$ in $E^3 - G$. To do this we consider a relation at each end point, branch point, or undercrossing point of G. A *relation* (designated by $r_i = 1$) is a word r such that it is permitted to cancel or insert either r_i or r_i^{-1} anywhere in any word. (If $r_i = x_1 x_2 \cdots x_s$, $r_i^{-1} = x_s^{-1} x_{s-1}^{-1} \cdots x_1^{-1}$.)

VI.5.A. *End point relation*. If A_j reaches out to v_i as shown in Figure VI.5.A, the loop f going under A_j is homotopic to the loop f' which agrees with f generally ex-

cept that it goes around the v_i end of A_j instead of crossing under. The relation $r_i = 1$ corresponding to v_i is $a_j = 1$.

VI.5.B. *Branch point relation*. If a loop f goes under the A's reaching out from a branch point as shown in Figure VI.5.B, it is homotopic to the loop f' partially shown. The relation $r_i = 1$ in this case is $a_s^{-1}a_k a_j a_r^{-1} = 1$.

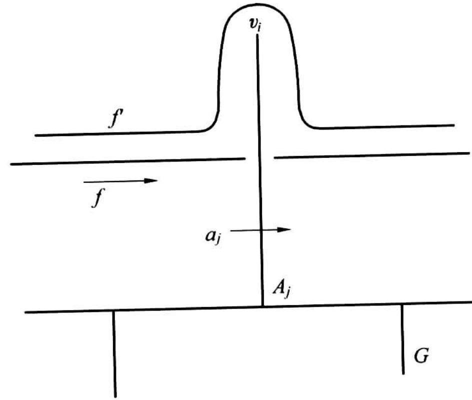


FIGURE VI.5.A

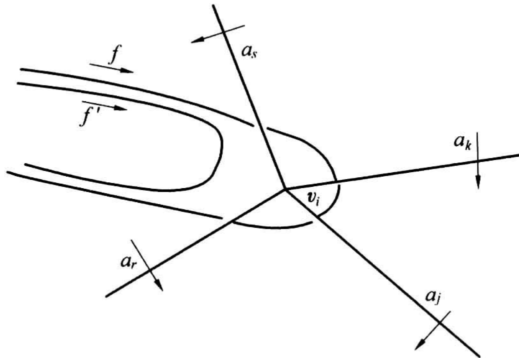


FIGURE VI.5.B

VI.5.C. *Undercrossing relation*. Suppose f goes around an undercrossing as shown in Figure VI.5.C Then $f' \sim f$. The relation $r_i = 1$ is $a_j a_s a_j^{-1} a_r^{-1} = 1$.

Suppose we associate a relation $r_i = 1$ with each end point, branch point, and undercrossing point of G. Then $f_1 \sim f_2$ in $E^3 - C$ if $W(f_1)$ can be changed to $W(f_2)$ by a finite sequence of the following operations:

(1) Inserting xx^{-1} or $x^{-1}x$ anywhere.

(2) Canceling xx^{-1} or $x^{-1}x$ anywhere.

(3) Inserting any r_i or its inverse anywhere.

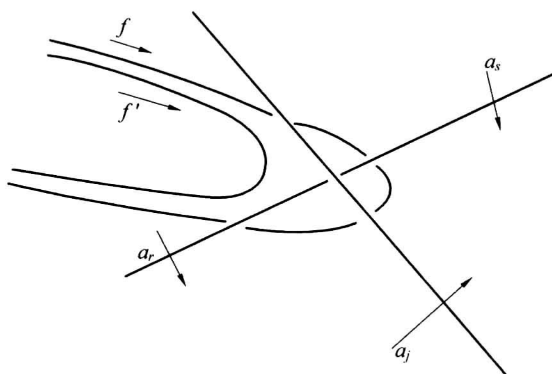


FIGURE VI.5.C

(4) Canceling any r_i or its inverse anywhere.

We shall show in the next section that the converse of the preceding paragraph is true in that if $f_1 \sim f_2$ in $E^3 - G$, then $W(f_1)$ can be changed to $W(f_2)$ with operations of these types.

We did not impose definite rules as to how to get the r_i's. For example, one person looking at Figure VI.5.C might have written $a_j a_s a_j^{-1} a_r^{-1} = 1$ and another might have written $a_s a_j^{-1} a_r^{-1} a_j = 1$. However, this does not matter since

$$a_s a_j^{-1} a_r^{-1} a_j = a_j^{-1} a_j a_s a_j^{-1} a_r^{-1} a_j = a_j^{-1} a_j = 1$$

To cancel out $a_s a_j^{-1} a_r^{-1} a_j$ one might first insert $a_j^{-1} a_j$ before a_s, then cancel out the middle four letters and then the remaining two. We do not need to choose any particular r_i while looking at Figures VI.5.B or VI.5.C. In fact, frequently, instead of reading a relation like $a_j a_s a_j^{-1} a_r^{-1} = 1$ from Figure VI.5.C, it might be simpler to use either $a_r = a_j a_s a_j^{-1}$ or $a_j a_s = a_r a_j$.

EXERCISE VI.5 Suppose $W(f_1) = a_1 a_3^{-1}$, $r_1 = a_1$ is an end point relation, and $r_2 = a_8 a_{10} a_7^{-1} a_{10}^{-1}$ is an undercrossing relation as in Figure VI.3.C. Describe an f_3 so that $f_3 \sim f_1$ and $W(f_3) = a_8 a_{10} a_7^{-1} a_{10}^{-1} a_3^{-1}$.

Solution. Slip the graph of f_1 from under a_1 to get an f_2 such that $f_1 \sim f_2$ and $W(f_2) = a_3^{-1}$. Next, slide the first part of the graph of f_2 so that it comes near the subscript 8 of a_8 in Figure VI.3.C. Push the curve down here and let it cross first under a_8 and a_{10} in direction of the arrows and then under a_7 and a_{10} in directions against arrows, and finally back up near the point of original descent and let if continue as before. There results an f_3 so that $f_2 \sim f_3$ and $W(f_3) = a_8 a_{10} a_7^{-1} a_{10}^{-1} a_3^{-1}$.

VI.6. Shelling. In this section we digress from considering fundamental groups to look at shelling. Suppose T is a triangulation of a PL disk D^2. A *shelling* of D^2 is an

ordering of the 2-simplexes $\Delta_1^2, \Delta_2^2, \cdots, \Delta_n^2$ of T such that if $1 \leqslant n_0 \leqslant n$, $\bigcup_{i=1}^{n_0} \Delta_i^2$ is a PL disk. We say that Δ_n^2 was shelled first and Δ_1^2 saved until last. This last Δ_1^2 is not shelled at all but left instead.

THEOREM VI.6.A. *Any triangulation of a 2-disk can be shelled so that any designated 2-simplex can be saved until last.*

PFOOF. The proof is by induction on the number of 2-simplexes in the triangulation T of a disk D^2. We are given Δ_1^2 and it is saved until last.

If some 1-simplex Δ^1 of T spans D^2, it separates D^2 into two disks D_1^2, D_2^2 so that $\Delta_1^2 \subset D_1$ and $D_1 \cap D_2 = \Delta^1$. Then shell D_2 saving until last the 2-simplex of T in D^2 containing Δ^1.

If no 1-simplex of T *spans* D^2, *start by using for* Δ_n^2 *any 2-simplex in* T *other than* Δ_1^2 *with an edge on Bd* D^2.

Instead of triangulating a disk we may subdivide it inot subdisks. It is to be understood that if two of the subdisks intersect, the intersection is a subset of the boundary of each. Such a subdivision is called a cellular subdivision of a disk. *A* shelling *of this subdivison is an ordering* $D_1, D_2, \cdots, D_n$ *of the subdisks such that each* $D_1 \cup D_2 \cup \cdots \cup D_i$ $(1 \leqslant i \leqslant n)$ *is a disk. The following is a variation of Theorem VI.6.*

THEOREM VI.6.B. Any cellular subdivision of a-disk can be shelled so that any-designated disk of the subdivision is saved until last.

Shelling off 3-simplexes in dimension 3 plays an important role in the study of 3-dimensional topology, but we postpone a discussion of this until the sixth section of Chapter XIV.

VI.7. Changing words. In this section we show that if $f_1 \sim f_2$ in $E^3 - G$ then $W(f_1)$ can be change to $W(f_2)$ with a finite number of operations of the following four types:

(1) Inserting xx^{-1} or $x^{-1}x$ anywhere.

(2) Canceling xx^{-1} or $x^{-1}x$ anywhere.

(3) Inserting any r_i or its inverse anywhere.

(4) Canceling any r_i or its inverse anywhere.

It simplifies matters to regard things as polyhedral so we suppose that G is a finite graph, P is a horizontal plane in regular position with respect to G, ρ is the vertical projection into P, $G' = \rho(G)$, and f_1, f_2 are piecewise linear loops normal with respect to G such that $f_1 \sim f_2$ in $R^3 - G$. No vertex of the graph of f_i is on $\rho^{-1}(G')$, and no point of the graph of f_i intersects any vertical line through either a vertex of G or an un-

dercrossing point of G.

Let S be a square disk in R^2 with opposite vertieces(0,0)and(1,1), h_1 be a PL homeomorphism of[0,1]onto the union of the left side and top of S, and g_2 be a PL homeomorphism of [0,1]into the union of the bottom and right side of S. Let F be a homeomorphism of Bd S into R^3 so that $F(h_1)=f_1$ and $F(g_2)=f_2$.

Since $f_1 \sim f_2$ in $R^3 - G$, there is a map $F': S \to R^3 - G$ such that $F' = F$ on Bd S. Let T be a triangulation of S of small mesh and F be a linear map of S under T into $R^3 - G$. At first we let $F = F'$ on vertices on T but then make needed shifts so as to get a linear map $F: S \to R^3 - G$ so that

$$F = F' \text{ on Bd } S$$

and for each 2-simplex Δ^2 of T, if $\rho F(\Delta^2)$intersects G', it intersects as shown in Figures VI.7.B,VI.7C,VI.7.D or VI.7E.

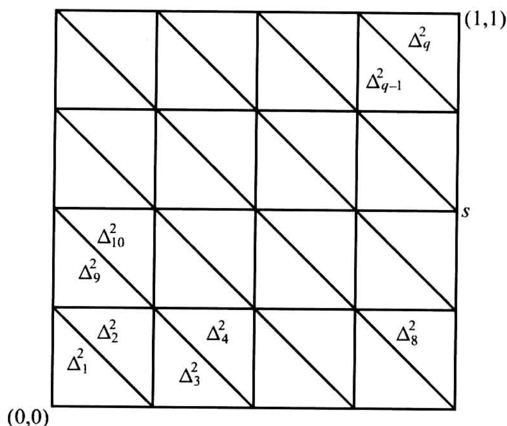


FIGURE VI.7.A

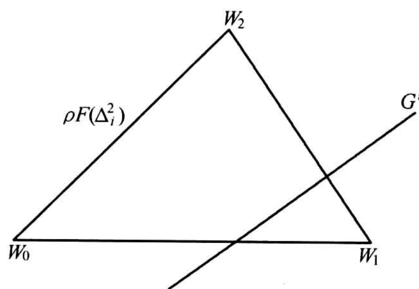


FIGURE VI.7.B

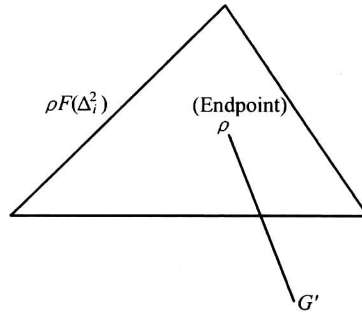


FIGURE VI.7.C

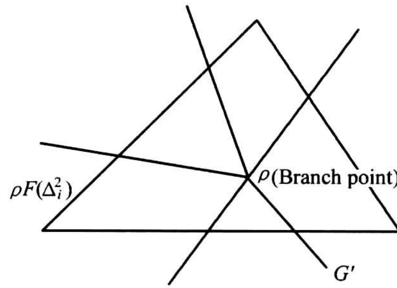


FIGURE VI.7.D

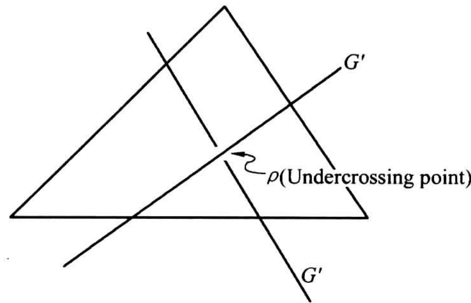


FIGURE VI.7.E

While we could cope with any shelling of any arbitrary triangulation of S, there is some simplicity in triangulating S as shown in Figure VI. 7. A and shelling off Δ_q^2, $\Delta_{q-1}^2, \cdots, \Delta_1^2$ in the order shown. Recall that h_1 took$[0,1]$onto the left side and top of S. We suppose Δ_q^2 is shelled and we get a new h(say h_2) that differs from h_1 in that $h_2[0,1]$ contains the slanted and vertical side of Δ_q^2 rather than the horizontal side. Also, $h_3(0,1)$contains vertical and horizontal sides of Δ_{q-1}^2 rather than a slanted side. As i increases, $h_i[0,1]$ moves down until $h_{q+1}[0,1]$ is the union of the bottom and right

side of S.

Note that $W(f_1) = W(F(h_1))$ and $W(f_2) = W(F(h_{q+1}))$. We complete verification that $W(f_1)$ can be changed to $W(f_2)$ by allowable operations by showing that for each i, $W(F(h_i))$ can be changed to $W(F(h_{i+1}))$ by such operations. We pass from h_i to h_{i+1} by shelling Δ_{q+1-i}.

If $\rho F(\Delta^2_{q+1-i}) \cap G' = 0$, $W(F(h_i)) = W(F(h_{i+1}))$ and no operation in necessary.

If $\rho F(\Delta^2_{q+1-i}) \cap G'$ resembles Figure VI.7.B, $F(\Delta^2_{q+i-1}) \cap G = \emptyset$ so G either passes below or above the triangle since $F(S) \subset R^3 - G$. If it passes below, $W(F(h_i)) = W(F(h_{i+1}))$, and if it passes below, the operation used depends on which sides of Δ^2_{q+1-i} belong to $h_i[0,1]$and $h_{i+1}[0,1]$. If one of the sides crossing under G belongs to $h_i[0,1]$ and the other to $h_{i+1}[0,1]$, $W(F(h_i)) = W(F(h_{i+1}))$ with no change. If both such sides belong to the same one of $h_i[0,1]$, $h_{i+1}[0,1]$, an operation of Type 1 or Type 2 changes $WF(h_i)$to $W(F(h_{i+1}))$.

If $\rho F(\Delta^2_{q+1-i}) \cap G'$ resembles Figure VI.7.C, G misses $F(\Delta^2_{q+i-1})$. If G is below, $W(F(h_i)) = W(F(h_{i+1}))$,and if it is above, an end point relation can be used to change $W(F(h_i))$ to $W(F(h_{i+1}))$.

Also, if Figure VI. 7. D shows the situation, G misses $F(\Delta^2_{q+i-1})$ and $W(F(h_i)) = W(F(h_{i+1}))$ or can be changed to it by using a branch point relation in cojunction with some operation of Types 1 and 2 to make the branch point operation have the proper start.

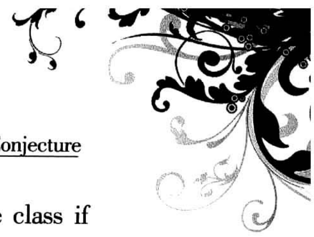
If Figure VI. 7. E describes the situation, there are more possibilities. If $F(\Delta^2_{q+i-1})$is above G, $W(F(h_i)) = W(F(h_{i+1}))$. If it is below, the procedure is the same as in consideration of Figure VI.7.D except that we use an undercrossing relation rather than a branch point relation. If one part of G is beneath $F(\Delta^2_{q+i-1})$and another part is above, we ignore the part below below and proceed as in the treatment of the case shown by Figure VI.7.B.

Suppose G is a PL finite graph in R^3 and f_1, f_2 are loops in $R^3 - G$ with a common base point. The principle results of § § VI.5 and VI.7 can be stated as follows.

THEOREM VI.7 *$f_1 \sim f_2$ in $R^3 - G$ if and only if $W(f_1)$ can be changed to $W(f_2)$ by a finite number or operations of the prescribed types.*

VI.3. Presentation of groups. We use $H = \{a_1, a_2, \ldots\}$ to denote the free group with generators $a_1, a_2, \ldots$. An element of H is an equivalence class of words (finite ordered set of letters)where the letters of the words are the generators $a_1, a_2, \ldots$





and their inverses $a^{-1}, a_2^{-1}, \ldots$. Two words belong to the same equivalence class if and only if one can be changed to the other with a finite number of operations where each relation consists of either the insertion of two adjacent letters xx^{-1} or $x^{-1}x$ somewhere in the word, or the cancellation of such a two-letter combination. The identity element is the class which contains the word with no letters.

Suppose $H' = \{a_1, a_2, \ldots \mid r_1 = r_2 = \cdots = 1\}$ is a group, and two words belong to the same equivalence class if one can be changed to the other with a finite number of operations of the following types:

(1) Inserting xx^{-1} or $x^{-1}x$ somewhere.

(2) Canceling xx^{-1} or $x^{-1}x$ somewhere.

(3) Inserting an r_i or its inverse somewhere.

(4) Canceling an r_i or its inverse somewhere.

An alternative way to look at $H' = \{a, a, \cdots \mid r_1, r_2, \cdots = 1\}$ is to let $H' = H/K$ where K is the smallest normal subgroup of H which has elements containing the r_i's.

We find from §§ VI.5 and VI.7 that $\pi_1(E^3 - G) = \{a_1, a_2, \cdots, a_n \mid r_1 = r_2 = \cdots = r_m = 1\}$ where the a_i's are the letters attached to the arrows going under arcs in G and the r_i's are the relations obtained at the end points, branch points, and undercrossing points of G.

If G_1, G_2 are two finite graphs in E^3 such that there is a homeomorphism h of E^3 onto E^3 such that $h(G_1) = G_2$, then $\pi_1(E^3 - G_1) = \pi_1(E^3 - G_2)$. However, one should not conclude the converse that if $\pi_1(E^3 - G_1) = \pi_1(E^3 - G_2)$ then there is such a homeomorphism h. Leaving a dangling feeler off a graph does not alter the fundamental group of its complement. In fact, there are smooth simple closed curves (tame knots) in E^3 such that the fundamental groups of their complements are the same even though there is no space homeomorphism taking one onto the other. However, the difference of the fundamental groups of their complements does imply the lack of a space homeomorphism.

There are rules for changing the presentation of a group but we shall not dwell on them here. We can add a generator if we add a relation to tell what the generator equals; a relation may be added that is a consequence of other relations; a relation may be deleted that is a consequence of other relations; if a relation tells what a generator equals, then the generator and the relation may be deleted if no other relation involves the generator.

VI.9. Short cuts. We mention several simplifictions that one discovers as he deals with the fundamental group of the complement of a graph in E^3. The first three involve

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

changint the graph and the last three involve simplifying the presentation of the group.

(a) *Simplifying projections*. If the regular projection of a graph G_1 looked like G'_1 as shown in the left half of Figure VI.9.a its complement would have the same fundamental group as the complement of graph G_2 obtained from G_1 by a homeomorpism of E^3 onto itself. The projection G'_2 of G_2 shown in the right half of Figure VI.9.a is simpler.

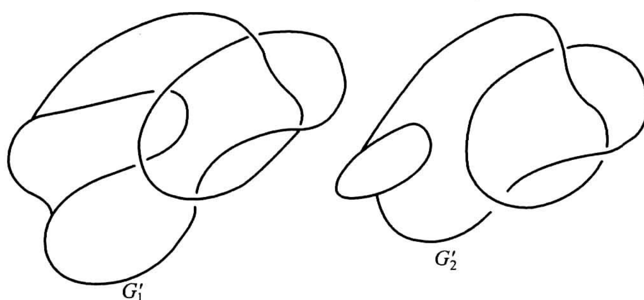


FIGURE VI.9.a

(b) *Shrinking arcs*. If a smooth arc in E^3 is shrunk to a point, the resulting decomposition space is topologically E^3. Let G_2 be a finite graph associated with the projection G'_2 shown on the right side of Figure VI.9.a. There is an arc in G_2 such that if this arc is shrunk to a point, G_2 becomes a finite graph G_3 whose projection is shown in the left side of Figure VI.9.b. Note that $E^3 - G_2$ is topologically equivalent to $E^3 - G_3$ and $\pi_1(E^3 - G_2) = \pi_1(E^3 - G_3)$.

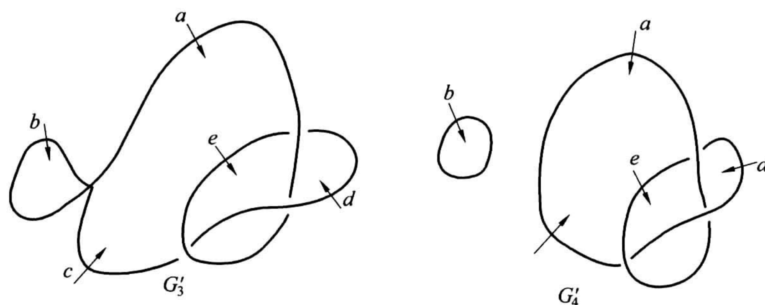


FIGURE VI.9.b

Shrinking arcs also enables us to eliminate dangling arcs from finite graphs. See Figure VI.5.A.

(c) *Pulling graphs apart*. Note that the relation $bb^{-1}ca^{-1} = 1$ obtained at the branch point of G_3 merely tells us that $a = c$. Since no other relation involves b, $\pi_1(E^3 - G_3) = \pi_1(E^3 - G_4)$, where the projection G'_4 of G_4 is shown in the right half



of Figure VI.9.b. In general, it is permissible to pull a graph G apart at a vertex v if the projection of v is a separating point of the regular projection G' of G. To compute $\pi_1(E^3 - G_1)$, it would be a short cut to compute $\pi_1(E^3 - G_4)$. One should not conclude that since $\pi_1(E^3 - G_3) = \pi_1(E^3 - G_4)$ that $E^3 - G_3$ and $E^3 - G_4$ are homeomorphic.

(d) *Droppinig a relation*. Instead of pulling a path f under a particular branch point or an undercrossing point as suggested in Figures VI.5.B or VI.5.C, it is possible to get an equivalent homotopy by pulling the path f to the path f' by moving under the rest of G. Hence any particular branch point or undercrossing relation r_i of $\pi_1(E^3 - G) = \{a_1, a_2, \cdots, a_n \mid r_1 = r_2 = \cdots = r_m = 1\}$ is a consequence of the others and there is no loss in dropping it from the presentation $\{a_1, a_2, \cdots, a_n \mid r_1 = r_2 = \cdots = r_m = 1\}$.

(e) *Equating generators*. Suppose G is a graph whose regular projection G' is shown in Figure VI.9.e. If one followed the suggestions made earlier, one would assign different letters to the vectors a_i as shown. However, one finds by either moving under the part of G on the left, or on the right, that a path going under the lower arc is homotopic to one going under the upper arc. If one writes down the presentation of $\pi_1(E^3 - G)$ as suggested in the preceding section, if follows from the algebra that $a_1 = a_2$. It is a short cut to designate each of a_1, a_2 by the same letter, say a, without going through the algebra.

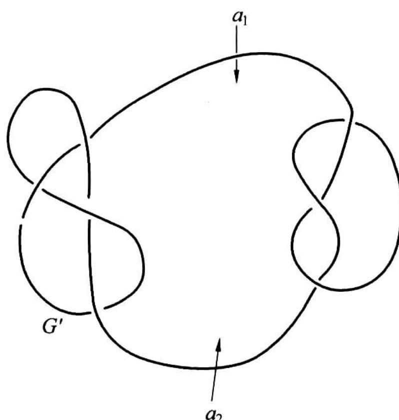


FIGURE VI.9.e

(f) *Further elimination of generators and relations*. Consider a graph G whose projection G' is shown in Figure VI.9.f. Instead of giving the vector c a name, it would be a short cut to call it $b^{-1}ab$. It is convenient not to assign separate letters to vectors associated with arcs of G that do not pass above points of G, since these can be as-

signed words instead of letters. In Figure VI.9.f we could drop c as a generator and combine the relations $c = b^{-1}ab$ and $c = d^{-1}ed$ to become the single relation $b^{-1}ab = d^{-1}ed$.

As a consequence of this observation, one obtains the following.

EXERCISE VI.9.A. Suppose G is a simple closed curve which is the union of n vertical spanning segments of a cube C and n arcs on the boundary of C. Show that $\pi_1(E^3 - G)$ has a presentation with n generators and $n-1$ relations.

Solution. In Figure VI.9.A we show the cube C with the front face smaller than the back and the curve G missing the back face. The generators associated with the arcs on the boundary of C are the only ones needed since others do not pass in front of other points of G. Hence $\pi_1(R^3 - G)$ can be written with n generators (and inverses). If one starts at the top of vertical segments and assigns words for each arc, there result n-relations. However, one relation is unnecessary.

EXERCISE VI.9.B. Suppose G is smooth simple closed curve with n undercrossing points. If $n \geqslant 2$, then $\pi_1(E^3 - G)$ has a presentation with $n-1$ generators and $n-2$ relations.

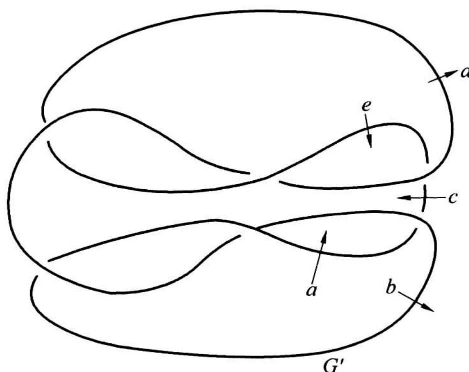


FIGURE VI.9.f

Solution. The n undercrossing points divide G into n arcs $A_1, A_2, \ldots, A_n$. Let $a_1, a_2, \ldots, a_n$ be generators associated with the A_i's.

If some A_i does not pass above an undercrossing point, the generator a_i can be written as the product of other a's, and the two relations involving a_i can be changed to a single relation not involving a_i. Also, we can ignore another relation so to end up using only $n-1$ generators and $n-2$ relations.

If each A_i passes above an undercrossing point, no one passes above two. If A_j passes above $\overline{A}_1 \cap \overline{A}_2$, we can replace a_1 by the appropriate three-letter word and drop

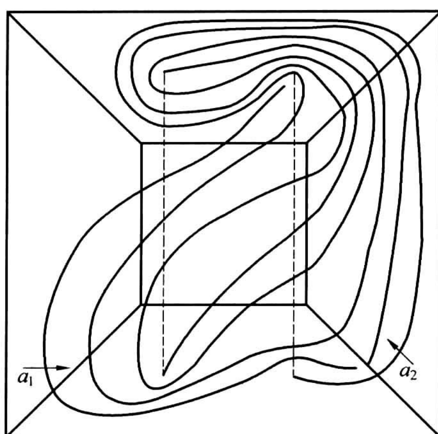


FIGURE VI.9.A

the corresponding relation. Dropping another relation gets us down to $n-1$ generators and $n-2$ relations.

EXERCISE VI.9.C. Can the polygon P shown in Figure VI.9.C be shrunk to a point in the complement of the simple closed curve J shown there?

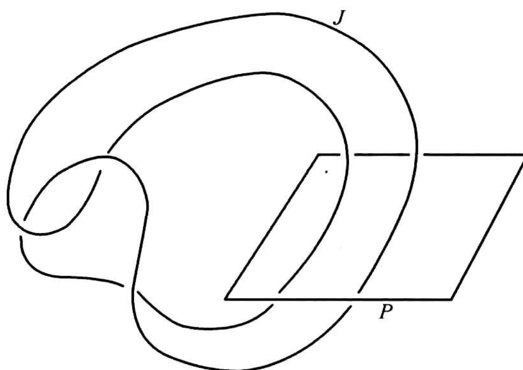


FIGURE VI.9.C

Answer. Yes. A computation of $\pi_1(R^3-J)$ shows that it is the infinite cyclic group and P corresponds to the identity element.

VI.10. Why compute fundamental groups? Some mathematicians study the fundamental groups of the complements of curves for their own sake and because they find the study interesting. We point out some applications of the study to the topology of 3-manifolds.

To show that the exterior of the Alexander horned sphere is not simply-connected, one might show that the simple closed curve J of Figure VI.10.A cannot be shrunk to a

point in $R^3 - X$.

EXERCISE VI.10.A. Compute and simplify $\pi_1(R^3 - X)$ of Figure VI.10.A.

Solution. $\pi_1(R^3 - X) = \{a, b, c, d, e \mid ecd^{-1} = eab^{-1} = bda^{-1}d^{-1} = ada^{-1}c^{-1} = 1\}$. Throw out e, b, c and the last relation to get $\pi_1(R^3 - X) = \{a, d\}$. Hence $\pi_1(R^3 - X)$ is a free group on 2-generators and J corresponds to the word $ada^{-1}d^{-1}$.

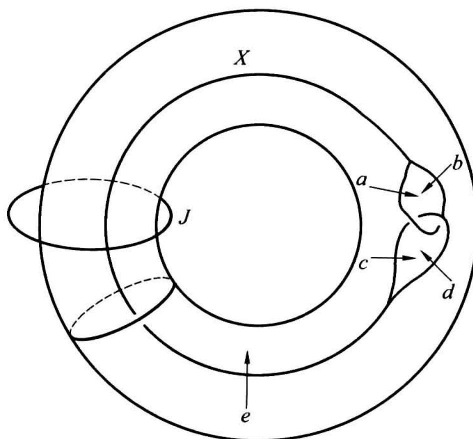


FIGURE VI.10.A

Sometimes the presentation for π_1 is so complicated that it is difficult to determine if the group is trivial. Even if $\pi_1(R^3 - G)$ is nontrivial, there is difficulty in deciding which loops in it are homotopic to constants. The loop theorem from Chapter XVII sometimes offers help. The proof of the following theorem illustrates this.

THEOREM VI.10.B *Suppose C is the cylinder shown in Figure* VI.10.B *with bases B_1, B_2; T_2, T_2 are two eye bolts embedded in C as shown so that T_1 and T_2 link while $T_i \cap \mathrm{Bd}\ C = B_i$; and X is a closed set in R^3 such that $X \cap C = B_1 \cup B_2$. If J is a simple closed curve in $R^3 - (X \cup C)$ that cannot be shrunk to a point in $R^3 - (X \cup C)$, then J cannot be shrunk to a point in $R^3 - (X \cup T_1 \cup T_2)$.*

PROOF. Assume that J can be shrunk to a point in $R^3 - (X \cup T_1 \cup T_2)$. Suppose J and Bd C are PL. Let T be a triangulation of Δ^2 and f be a linear map of(Δ^2, T) into $R^3 - (X \cup T_1 \cup T_2)$ such that f takes Bd Δ^2 homeomorphically onto J, f is a homeomorphism on each simplex Δ of T, and each $f(\Delta)$ is GP with respect to Bd C. Each component of $f^{-1}(\mathrm{Bd}\ C)$ is a simple closed curve on Δ^2. Induction on the number of components of $f^{-1}(\mathrm{Bd}\ C)$ is used to show that f does not exist. Assume $f^{-1}(\mathrm{Bd}\ C)$ has k components and Theorem VI.10.B is true if it has fewer. Let J_1 be a



simple closed curve of $f^{-1}(\text{Bd } C)$ that bounds an innermost disk D_1 in Δ^2. Then f on J_1 is not homotopic to a point on Bd $C-(B_1 \cup B_2)$; otherwise we could redefine f near D_1 and eliminate a component of $f^{-1}(\text{Bd } C)$.

Although J_1 is a simple closed curve, $f(J_1)$ may be singular. However, it follows from the loop theorem (see Chapter XVII) that there is a simple closed curve K on Bd $C-(B_1 \cup B_2)$ such that K does not bound a disk on Bd $C-(B_1 \cup B_2)$ but K does bound a disk E in the one of $C-(T_1 \cup T_2)$ or $R^3-(X \cup \text{Int } C)$ containing $f(D_1)$. We show that this is impossible.

The simple closed K on Bd C separates B_1 from B_2 since it does not bound a disk on Bd $C-(B_1 \cup B_2)$. It follows from Exercise VI.10.A that $E \not\subset C-(T_1 \cup T_2)$.

If $E \subset R^3-(X \cup \text{Int } C)$, then J can be shrunk to a point in $R^3-(X \cup \text{Int } C)$. The assumption that there was an f led to the contradiction that there is no place for E.

THEOREM VI.10.C. *The complement of the solid Alexander horned sphere is not simply-connected.*

PROOF. Exercise VI.10.A shows that the simple closed curve P of Figure IV.3.C cannot be shrunk to a point in R^3-H_1. It follows from repeated applications of Theorem VI.10.B that is cannot be shrunk to a point in any R^3-H_i. The compactness of the image of Δ^2 under a map shows that it cannot be shrunk to a point in the complement of $R^3-\bigcap H_i$.

Suppose X of Figure VI.10.D is a modification of the Fox-Artin arc as shown in Figure IV.9.A where a solid cone C_0 to the left of F_1 and a solid cone C_1 to the right of F_1 are added. To help compute $\pi_1(R^3-X)$ we have substituted dotted triods for the cones.

EXERCISE VI.10.D. Compute $\pi_1(R^3-X)$.

Answer. $\pi_1(R^3-X)$ is a free group on two generators.

The following result follows from an algebraic computation of $\pi_1(R^3-X)$.

EXERCISE VI.10.E. The simple closed curve J_1 shown in Figure VI.10.D cannot be shrunk to a point in R^3-X.

We are now ready to prove that the complement of the Fox-Artin arc of Figure 9.A is not simply-connected. Let A denote the arc and C_0, C_1, C_2 be the cones shown in Figure VI.10.F.

EXERCISE VI.10.F. J_1 cannot be shrunk to a point in $R^3-(A \cup C_0 \cup C_2)$.

PROOF. Following the proof of Theorem VI.10.B, we find that the assumption that J_1 can be shrunk to a point in $R^3-(A \cup C_0 \cup C_2)$ leads to a disk E in $R^3-(A \cup C_0 \cup C_2)$ such that $E \cap \text{Bd } C_1 = \text{Bd } E$ and Bd E does not bound a disk on Bd $C_1 - A =$ Bd

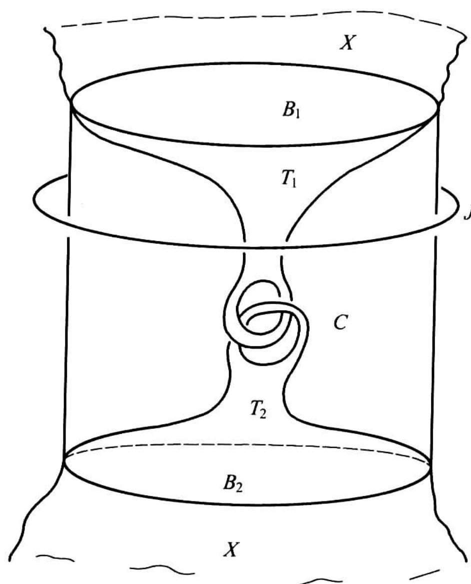


FIGURE VI.10.B

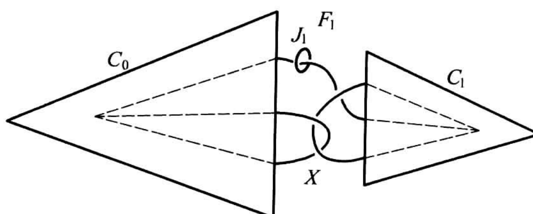


FIGURE VI.10.D

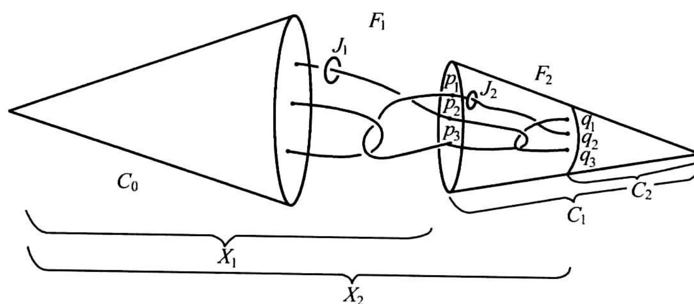


FIGURE VI.10.F

$C_1 - \{p_1, p_2, p_3\}$. Again we find that E cannot be in C_1 or in $R^3 - (A \cup C_0 \cup \text{Int } C_1)$.

We proved Exercise VI.10.F by showing there is no place for E.



The following result can be proved from repeated application of Exercise VI. 10. F.

THEOREM VI. 10. G. *The complement of the Fox-Artin arc shown in Figure* IV. 9. A *is not simply-connected.*

The following two exercises are related to Antoine's necklace.

EXERCISE VI. 10. H. The simple closed curve J shown in Figure VI. 10. H cannot be shrunk to a point in the complement of $T_1 \cup T_2$ shown there.

Solution. $\pi_1(R^3 - (T_1 \cup T_2)) = \{a, b\}$ and J corresponds to the word $bab^{-1}a^{-1}$.

EXERCISE VI. 10. I. If there are three or more T_i's in FIGURE VI. 10. I, then J cannot be shrunk to a point in $R^3 - \bigcup T_i$.

PROOF. Let M_1 be a plug with two holes filling T_1 as shown. First we show that if J can be shrunk to a point in $R^3 - (\bigcup T_i)$, it can be shrunk in $(R^3 - \bigcup T_i) - M_1$.

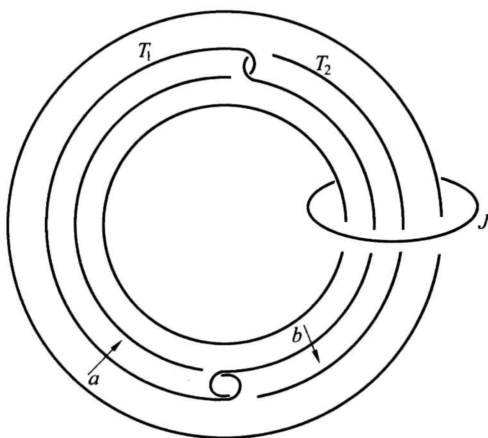


FIGURE VI. 10. H

If J can be shrunk in $R^3 - \bigcup T_i$ but not in$(R^3 - \bigcup T_i) - M_1$, it follows from the argument used in the proof of Theorem VI. 10. B that there is a disk E in $M_1 - \bigcup T_i$ such that $E \cap \text{Bd } M_1 = \text{Bd } E \subset \text{Bd } M_1 - (\bigcup T_i)$, and Bd E does not bound on Bd $M_1 - (\bigcup T_i)$. Then one or both of T_2, T_n prevent D_i from lying either in M_1 or $R^3 - M_1$. Hence J can be shrunk to a point$(R^3 - \bigcup T_i) - M_1$.

Let M_2 be a plug with one hole which finishes finishes filling T_2. An argument similar to that used above shows if J can be shrunk in $R^3 - \bigcup T_i$ it can be shrunk in $(R^3 - \bigcup T_i) - (M_1 \cup M_2)$. By putting in more plugs we find that if J can be shumrk to a point in $R^3 - \bigcup T_i$, it can be shrunk to a point in the complement of a center line of the big torus T. This, of course, is impossible.

EXERCISE VI. 10. J. Can the polygon P of Figure VI. 10. J be shrunk to a point

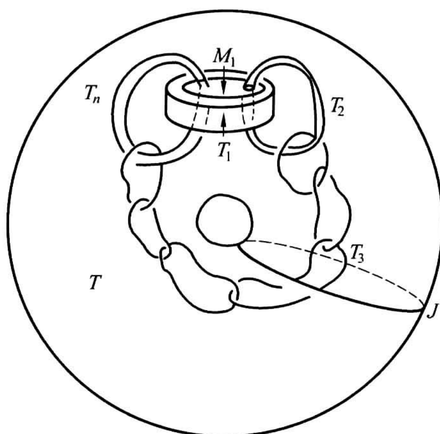


FIGURE VI.10.I

in $R^3-(T_{11}\cup T_{12})$?

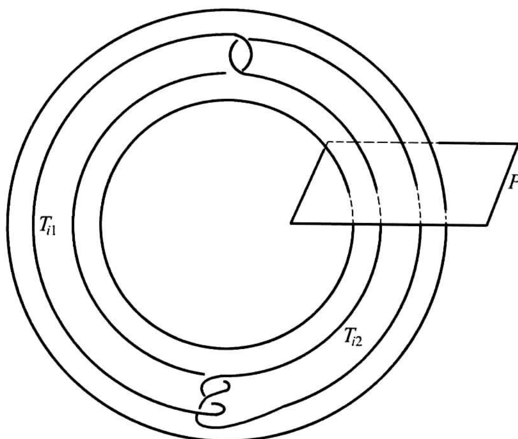


FIGURE VI.10.J

Answer. Yes. A computation of $\pi_1(R^3-(T_{11}\cup T_{12}))$ shows that it is a free abelian group on 2-generators and P corresponds to the identity element.

It is an unsolved question (the *Poincaré conjecture*) as to whether or not each compact connected 3-manifold whose fundamental group is trivial is a topological 3-sphere. In trying to construct counterexamples, one might remove from S^3 a tubular neighborhood of a smooth finite graph in S^3 and sew this tubular neighborhood back in differently. It is known that if the Poincaré conjecture is false, a counterexample can be obtained in this way. The fundamental group of the complement of the tubular neighborhood is the same as the fundamental group of the complement of the graph. It is not known if the study of fundamental groups will be of any help in solving the Poincaré conjecture, but hops that





it may have encouraged an interest in these groups.

There is a distinction between being able to shrink a curve to a point and the curve bounding a disk. Algebra is sometimes of help in deciding whether or not a map is homotopic to a constant. For example, it shows that P of Figure VI.9.C shrinks in $R^3 - J$, but it does not tell us if it bounds a disk there. It does not tell us if P bounds a singular disk in $R^3 - J$ whose set of singularities is topologically starlike. We seem to need stronger tools to decide the Poincaré conjecture. Perhaps there is a 3-manifold whose fundamental group is presented in some complicated exotic fashion and that by some clever trick the group can be shown to be trivial. If the clever trick did not cause simple closed curves to lie in 3-cells in the 3-manifold, we would have a counterexample to the Poincaré conjecture.

VI.11. A homology cube. The cube with a knotted hole shows in Figure VI.11 is denoted by M^3. A center line for the knotted hole is provided with short oriented normal vectors a, b, c to facilitate the computation of $\pi_1(M^3)$ by methods we have used earlier in this chapter. One finds that $\pi_1(M^3) = \{a, b, c \mid aba^{-1}c^{-1} = bca^{-1}c^{-1} = 1\}$. While this presentation can be simplified, it is not our purpose to do so.

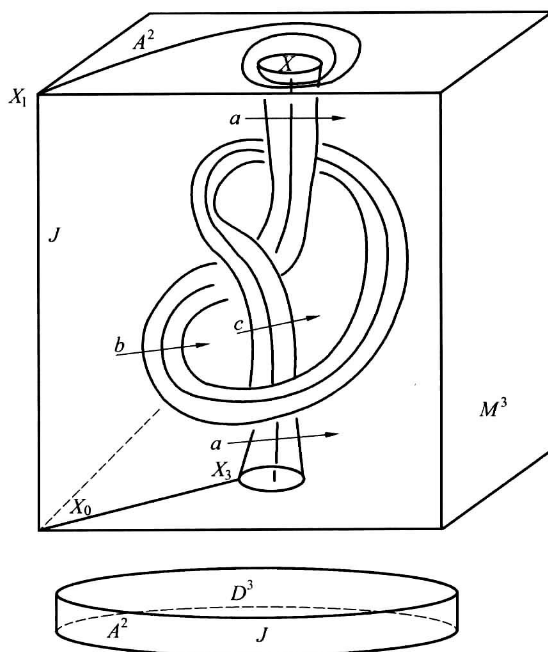


FIGURE VI.11

Let J be the simple closed curve on Bd M^3 associated with the loop which rises

from x_0 to x_1, goes from x_1 to x_2 on the top while circling the hole twice as shown, goes down the side of the hole to x_3, and returns to x_0 on the bottom. Then J corresponds to the word $a^2c^{-1}a^{-1}b^{-1}$.

Let P^3 be the pill box $D^2 \times I$ shown in the lower part of Figure VI.11, where its lateral side is denoted by A_2^2. § XVII.1 has a definition of a pillbox. Use A_1^2 to denote an annulus on Bd M^3 whose center line is J and let h be a homeomorphism of A_1^2 onto A_2^2. Suppose M^3 is sewed to P^3 along A_1^2 and A_2^2 by h as described in IV.2.B and IV.2.D. Then $K^3 = M^3 \cup_h P^3$ has a 2-sphere as a boundary and $\pi_1(K^3) = G = \{a, b, c \mid aba^{-1}c^{-1} = bca^{-1}c^{-1} = a^2c^{-1}a^{-1}b^{-1}\}$. One may note that if G is abelianized by adding relations $ab = ba$, $ac = ca$, $bc = cb$, then the first two relations imply $a = b = c$ and the last implies $a = 1$. Hence, the abelianization of $\pi_1(K_3)$ is trivial.

Sometimes K^3 is called a *homology cube*. One could get many other examples of such objects by boring different holes in cubes, but then x_1x_2 might need to circle the hole a different number of times to make the abelianization of the fundamental group trivial. Some require that the homology cube be topologically different from a 3-cell but we do not.

Question. Is there some nontrivial knotted hole in a cube such that if a pill box is added as described, then the fundamental group of the union is trivial?

VI.12. Other treatments. The techniques we have used for computing the fundamental group of the complement of graphs might be called the shelling method after our treatment of Theorem VI.7. This treatment[**B₂₃**] is covered in a Technical Report issued by Washington State University in 1965 by R. H. Bing entitled "Computing the Fundamental Group of the Complement of Curves". There is another widely used method where we enlarge G by digging more tunnels so that the fundamental group of the complement of the enlarged graph is free abelian. The extra tubes are then plugged and a theorem by van Kampen is used to compute $\pi_1(R^3 - G)$. Treatments of this technique are given by Crowell and Fox[**CF**], Fox[**F₃**], and Rolfsen[**R₁**].





附录X How Famous Can a Function Theorist Be①

The Twenty-Second Problem

Koebe and Poincaré

When Paul Koebe traveled he registered in hotels under assumed names, lest the waiters and chambermaids besiege him with questions about whether he were related to the famous function theorist, Koebe. If they were to find out that he was the great man himself, he would have no peace. It is safe to say that few mathematicians have had to be so careful. Hans Freudenthal, whose article in *The Mathematical Intelligencer* is the source of much of the biographical information in this short section, wrote:

When the first day of my study I was inscribed at the Mathematical Institute of Berlin University, it happened that Bieberbach was around and heard that I came from Luckenwalde. He turned and asked me: "So you are one of the Luckenwalde streetboys who run after Koebe to call 'there goes the famous function theoretician,' are you?"

Koebe did hail from Luckenwalde, where his family, in Freudenthal's words, "ran a renowned fire engine plant." One wonders if something in the Water of Luckenwalde encouraged native sons to talk of "renowned" fire engine plants (or function theorists). Luckenwalde provided Koebe many advantages. Born in 1882, he enjoyed the prosperity of his family and received and excellent education, attending gymnasium in Berlin and then continuing at the University of Kiel in 1900. He finished at Berlin University in 1905. He also attended Charlottenburg Technische Hochschule, a more practically oriented school, perhaps with an idea of picking up useful skills for running a fire engine plant. He habilitated at Göttingen in 1907, which placed him in a setting where many mathematicians were thinking about uniformization theory and Hilbert's twenty-second problem. In 1907 Koebe, independently of Poincaré, solved Hilbert's twenty-second problem and achived a measure of genuine fame—at least within the mathematical community. He then continued his work as a function theorist with single-minded determination and never strayed, doing other important work in this area. Koebe remained at Göttingen until 1910, when he moved on to Leipzig as extraordinarius. Thence to Jena

① 转载自 Ben H. Yandell. The Honore Class—Hilbert's Problems and Their Solvers.

in 1914 as full professor and back to Leipzig in 1926.

Koebe was respected, due to his talents, but considered "conceited and disagreeable" at Göttingen. He had a reputation for stealing the ideas of younger mathematicians. Reid quotes Courant to the effect that Koebe stole the idea of his thesis, not yet published, of which the two had spoken. Koebe rushed off a paper on the same subject and thus claimed independent credit for the result. Then Koebe showed up at a seminar where Courant was scheduled to speak on his thesis and by virtue of seniority took the position of speaking first. This was not well received by Courant's friends, according to Reid. They rigged up an elaborate apparatus and hid it in a chamber pot under the lectern of a class Koebe was teaching. An alarm sounded at erratic intervals as Koebe lectured. When he finally pulled the mechanism out of the chamber pot, it was greeted with much laughter. One of the mischief makers, Kurt Hahn, saw to it that an article describing the incident appeared in the local paper.

Freudenthal says that Koebe's eccentricities were so widely known that "younger people, when introduced to Koebe, almost automatically reacted: 'Ah, the great function theoretician,'" Such a greeting, instead of being recognized for what it was, succeeded in getting an assistantship for a friend of Freudenthal's.

Freudenthal also syas an "unidentified" man wearing a big hat, a turned-up collar, and dark blue glasses gained access to a Göttingen printer and inserted a footnote into a paper of Brouwer's on uniformization, acknowledging priority to Koebe. This incident supposedly happened in 1912, when Koebe had already moved on to Leipzig, so it contains the additional oddity of a possible special trip. Freudenthal does not vouch for the truth of this story and adds that later Koebe said it was a practical joke played on him. At the least, it captures something about how Koebe was perceived in Göttingen.

Freudenthal also wrote the entry on Koebe in the *Dictionary of Scientific Biography*, which ends:

Koebe's mathematical style is prolix, pompous, and chaotic. He tended to deal broadly with special cases of a general theory by a variety of methods so that it is difficult to give a representative selective bibliography. Koebe's life-style was the same: Koebe anecdotes were widespread in interbellum Germany. He never married.

Koebe died Auust 6, 1945.

A national hero of France and a best-selling author of exphanations of mathematics and science to the layman, Jules Henri Poincaré really was famous. By some, he was considered an important philosopher. He was admitted to the French Academy, a body of forty writers that passes on the correctness of words and usage in the French language,



despite the claim that he never revised a sentence. Poincaré was the greatest mathematician of his time, his only competition being Hilbert. Mathematicians compare him to Gauss, which is the highest honor. Blink at his solution of the twenty-second problem in the context of his whole mathematical career and you will miss it, and it was he who first asked the question.

Poincaré published around 500 papers in his lifetime, counting only papers reporting original research. His Oeuvres are eleven large volumes. each labeled a tome. While the first appeared in 1916, four years after Poincaré's death, the last did not come out until 1954, so monumental was the labor involved in selecting what was to be published. A separate three-volume set on celestial mechanics is important. Poincaré lectured on a different topic in mathematical physics every year at the Sorbonne and other institutions, like the École Polytechnique, and auditors turned many of these lectures into books.

Though prolific, Poincaré was a deep mathematician. Just as it took decades for the mathematical community to begin to understand and use the ideas contained in Riemann's eight-page paper on the zeta function, mathematicians find the same kind of depth in Poincaré again and again. Insight appear that are clearly expressed for their time, yet for which a precise mathematical vocabulary and technique would not be available for decades.

Poincaré came from a distinguished French family full of professionals, often of a technical type, and had a highly intelligent, devoted mother. Nearly all of my account of his early years comes from Gaston Darboux's "*Éloge Historique D'Henri Poincaré*", which appeared as a long obituary in *Mémoires de L'Académie des Sciences de L'Institut de France* in 1914. It also appears at the beginning of the second volume of Poincaré's *Oeuvres*. Darboux was the French mathematician whom Hilbert had the temerity to eulogize during the First World War. He had known Poincaré a long time, having been one of the examiners for his doctoral degree and for the École Normale. His account has a tone of Olympian praise, but I believe it was written with care and is accurate.

Poincaré is often mentioned and there are many books about aspects of his work, but astoundingly, nearly ninety years after his death, there is no comprehensive, modern scientific biography of him in English or any Western European language that I have been able to find. E.T. Bell wrote in 1937: "To give an adequate idea of this immense labor one would have to be a secound Poincaré, so we shall presently select two or three of his most celebrated works for a brief description." This echoes H. F. Baker in his 1915 obituary for the Royal Society of London: "To give any complete account of his

work is a task well nigh impossible on account of its vast range."

Poincaré's parents were "*lorrains tous les deux*," from the Alsace-Lorraine region in the northeast corner of France. Poincaré's father's family was originally from Neufchâteau. An ancestor who died in 1750 was an official of the court, *conseiller au bailliage de Neufchâteau*. Another member of the family was a mathematician at the *collège de Bourmaont* near Neufchâteau. Poincaré's grandfather, Jules Nicholas, a pharmacist, was born in Neufchâteau in 1794. He was stationed at the San Quentin military hospital in 1814 during the Napoleonic Wars. In 1817, at twenty-three, he moved to Nancy, taking along his parents and sisters. The family acquired a distinguished house, which sat between the ducal palace and the *porte de la Craffe*. When Poincaré was accepted into the French Academy, a description of this house was included in the welcoming speech by Frederic Masson: "solid, massive and without ornament."

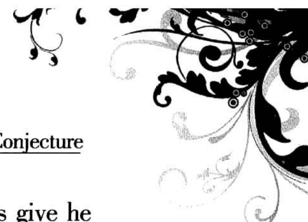
Into this house was born Leon Poincaré, Henri's father, in 1828. Leon became a prominent physician and professor of medicine in Nancy. Darboux writes that the Academy of Sciences received his reports with high appreciation and describes him as having a spirit "*très original*." Leon's younger brother Antoni, born in 1829, achieved a brilliant rank at the École Polytechnique and sent many communications to the French Academy about meteorology. He became inspector general of bridges and embankments. Henri's first cousin Raymond Poincaré, one of Antoni's sons, was president of the French Republic during the First World War and four times premier. Another son of Antoni's, Lucien, became "*directeur de l'En-seignement secondaire au Ministère de l'Instruction publique et des Beaux-Arts*." Darboux lists other distinguished Poincarés and in a footnote even traces the name (possibly) back to 1403 and a Petrus Pugniquadrati, a student at the University of Paris.

Darboux does not list an equivalently rich family tree for Poincaré's mother. In fact, as Bell points out, he neglects even to mention the family name (Launois), at least when she is first introduced, but he does mention her family's country estate in Arrancy, where Henri spent some of the best times of his childhood. Poincaré's mother is described as "*très bonne, très active et très intelligente*." Gaston Julia writes in *Le Livre du Centenaire de la Naissance de Henri Poincaré* 1854 ~ 1954 that Poincaré's maternal grandmother had a gift for mathematics.

Henri was born on April 29, 1854, into this lively and intellecctual family. He had one younger sister. Her son Pierre Boutroux (1880—1922) became a mathematician.

Darboux says that Poincaré had an exceptionally happy, even blissful, childhood. "His mother watched over him with a solicitude full of intelligence." He spoke very





young and had trouble at first expressing the rush of his thoughts. When he was give he caught diphtheria. Today we have largely forgotten what this illness entails. It starts as a mild sore throat but progresses rapidly as a false membrane forms on the surface of the throat and toxins poison the body. Death often came rapidly. Poincaré suffered from a paralysis of the larynx that lasted for nine months. The illness left him weak and timid for a long time. He wouldn't even dare to go down stairs by himself and couldn't stand the rough play of children his own age. During this period he became an avid reader. He never read the same book twice, and at this point the power of his memory became evient. Darboux says that he could always remember where in a book, on which page and on which lines, he had read something. Though perhaps exaggerated, the spatial aspect rings ture. We hear that Poincaré was particularly taken by a book of natural history, *La Terre avant le déluge*, by Louis Figuier. This eraly bestseller of science is a provocatively written and gorgeous book. Beautifully colored maps of the continents in surprising arrangements and shapes during different paleological periolds apear throughout.

Poincaré's schooling began with a tutor named Hinzelin who had published well-respected educational books and who opened up many subjects to young Henri without requiring much written work. Poincaré's mother was nervous about this when Henri entered the lycée at the age of eight, but he received the highest grade on his first essay. With a few notable exceptions he would always be first in school. A notebook survives from this first year of school that Darboux says makes clear Poincaré's exceptional talents. His work on history and geography stands out, and a French composition is particularly noted as a "little masterpiece." The exceptionally poor quality of his handwriting is also remarked upon.

School did not seriously inconvenience Poincaré. Homework got done in a haphazard though effective way. From a General(Paul)Xardel, who grew up with Henri(Darboux refers to him only as General Xardel throughout), we have the image of boks and homwork on a table, with Poincaré floating around the room, in and out, talking when there was conversation available, answering questions about the homework when asked. When he went to write he would pick up the pen with whatever hand was available.

Instead, the future member of the Academy liked to play. Since he was neither strong not coordinated, he preferred games for which imagination and spirit were more important. Sometimes he played with the sisters of his friends. During summer vacations with his maternal grandparents at Arrancy, Henri was allowed freedom to wander in the large garden or the countryside. He liked to walk rapidly through the garden with a

walking stick at his side. Sometimes he would stop and draw with the stick in the sand. He particularly loved animals and had a chance to observe them in the garden. Once he accidentally shot a bird in a tree and was greatly saddened. After that he would never shoot a gun.

Darboux reports that Poincaré was a loving son and brother. He was modest and didn't try to prove his superiority over others. However, when he had good reasons he could be immovable in his "passive resistance." Darboux quotes General Xardel, narrating a vacation the two families spent together in 1865: "Henri wanted to see everything, understand everything and explain to us everything." At the famous Echo of Ramberchamp Henri delivered a complete theory of the echo. Xardel said of Poincaré, "He had the gaiety and the expansiveness of a child, the reason of a man."

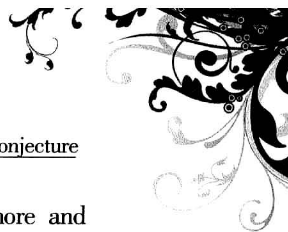
When he was thirteen, Poincaré went to the Exposition in Paris and emerged with ideas about politics. This led to the founding of a republic called the "Trinasie" in the gardens of Arrancy, with his sister and a cousin. It was a federation of three parts, one for each child, and each had their own language as well as a common language, which they used to conduct the government. Poincaré found that without seeming to he could effectively take all the power, but after discovering this he never did it again. The Trinasie lasted a number of years and was the trellis upon which many games grew.

Much of the Summer would pass with this free play in the garden with just one or two other children. But then there would be a week of visiting and parties with relatives in the region. Poincaré participated actively in the amateur theatricals and charades. When he was thirteen or fourteen he wrote a verse drama on Joan of Arc—also a *lorraine*. Poincaré liked to dance; Darboux says he never tired. This active social life continued into the school year and his studies never suffered.

During this period of play Poincaré was demonstating the ability to concentrate when something grabbed his attention—a game, a book, an idea. When interested he would not let go, even if the time set aside for the subject was over. This focus caused him to sometimes forget to pay attention to the world outside. He never could remember if he had eaten. Later in life Poincaré described himself as having a poor memory, and this is a way to understand that evaluation from a man who remembered everything he read. Darboux tells of a time when Poincaré was seven or eight and was walking with his mother and sister, lost in thought, when he discovered he was on the other side of a small river from them. He remedied the situation by wading in a straight line through the river up to his waist to get back on the correct path.

When he was fifteen Poincaré discoverd his passion for mathematics. Darboux





writes, "From this moment this vocation did nothing but grow and become more and more dominating and absorbing." Unlike the extreme case of Ramanujan, though, it did not keep him from pursuing his other studies. David Ruelle says that Kolmogorov had a theory, offered as an explanaton for the childish behavior of a colleague, that psychological maturation stopped the moment mathematical talent emerged. By this account Poincaré was allowed to mature to a relatively advanced mental age.

Poincaré passed his first set of exams with distinction, acquiring a *baccalauréat* in letters first and following it three months later with a degree in science. His Latin exam was evn better than his French exam. Oddly, he did poorly on the written exam in mathematics. He arrived late and appeared to not understand the question, a simple one on a convergent geometric series. Poincaré did very well on the oral examination and the examiners knew of his reputation. One of the examiners said that if any other candidate had presented such an exam he would have failed. The *baccalauréat* degress, earned at the age of seventeen, marked the end of what Americans call high school. However, for students headed for the *Grandes Écoles*, the most elite schools in France, two more years were spent at the lycée. Poincaré passed on to a course called"elementary mathematics."

The Franco-Prussian War, fought in great part in the northeast corner of France where Nancy lies, disturbed Poincaré's years in lycée. The French army attacked the Prussian army. The Prussians had been studying the lessons of the American Civil War and understood that with the new generation of firearms the defending army could sit back and slaughter the advancing forces. The French were slaughtered and surrounded. The French regime fell and the Commune of Paris(1871)had its day. Then the Germans took Paris and the weak and vacillating Third Republic was installed. The Prussians annexed the Alsace and part of the Lorraine, but luckily for Poincaré, Nancy remained as part of France, though closer to the border than it had been. The occupation was not polite. Poincaré had been too young and delicate to serve in the army but had accompanied his father in an ambulance to tend the wounded. He waw the brutality of war close up, but saw more personally compelling destruction when he accompanied his mother and sister to Arrancy. The battle of Saint-Privat had been fought nearby. On the journey they were forced to pass "in glacial cold, through burnt villages, empty of inhabitants". They found the family home where they had spent so many summers wrecked and looted. Poincaré's grandparents were fed on the day of the looting by a poor villager. In an interesting side note, the Japanese, who had been studying military strategy under the French, switched to Germany, which is partly why Takagi ended up in Göttingen rather

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

than Paris.

It is said that these events made Poincaré a lifelong patriot. They certainly had an effect on the attitude of his more political cousin, Raymond. During the period of occupation Henri learned German so he could read the only newspapers available. Despite the poor behavior of the German soldiers, Poincaré went out of his way throughout his career to apportion credit fairly to German colleagues. In some of his first major work he called what are now more commonly referred to as automorphic functions, at least in English, Fuchsian functions, after the German mathematician Lazarus Fuchs, who had some early involvement with them. In the same context he referred to Kleinian groups, since Felix Klein was working on the subject at the same time as Poincaré. When Poincaré delivered a series of six lectures in Göttingen in 1909 at Hilbert's invitation, hedelivered five in German (although some of the Göttingers were a bit miffed that he had the nerve to talk about integral equations, about which Hilbert was in the process of making substantial advances).

France and Poincaré continued with remarkable continuity after the interruption of 1870 ~ 1871. In 1872, Poincaré competed in a *concours général* given to all lycée students in France and won first prize. The next year he took the "special mathematics" course and again took first prize. During this last year at the lycée, Poincaré met two young mathematicians, Paul Appell (1855—1930) and Colson (probably Clement, 1853—1939), who also became prominent. Appell had moved to Nancy to escape the Germans, as his home town of Strasbourg had become part of the German Empire. On the first day of class Appell and Colson sized up Poincaré. They were perched above him in the tiered classroom and could observe the notes Poincaré was taking, or not taking. He was using a funeral announcement for paper. The notice appeared day after day and occasionally Poincaré would scrawl(*griffonner*)a line or two on it. They were astounded. They thought perhaps he wasn't serious. However, he had won first prize in all of France the previous year in the *concours*. They approached a more advanced student(the lycée in these final two years flowed into or blended with École Forestiere in Nancy)and persuaded him to ask Poincaré a difficult and "particularly obscure" question. Poincaré gave the answer"without thinking a minute."

Appell described Poincaré as a student:

From the beginning of term his intelligence revealed itself immediately to our professor Elliot as it did to us schoolfellows; he had the genial gift of immediate comprehension of the particular detail of each equestion, the general idea form which it came, and the place it occupied in the whole. He also had this implicity, this horror of effect, this





good sence of someone from Lorraine, this sureness in friendship that he retained all his life.

Poincaré took the exams for the École Polytechnique and the École Normale, the two most elite schools in France—doorways to the highest positions in French society, including the civil service. There is a bewildering array of institutions of higher learning in Paris, and many scholars are on the faculty of more than one. ph. D. s are granted by the universities. Poincaré chose the École Polytechnique, partly as a family tradition, entering first in his class. He did more poorly in the exam for the École Normale, finishing only fifth(in France). Mathematics occupies a particularly high place in these exams. Darboux regarded the fifth-place finish as a small failure. Appell finished second and chose the École Normale.

Darboux, writing forty years later, was able to find many people who remembered specific details of Poincaré's performance in his school exams. In one story, Poincaré spoke slowly, then closed his eyes. He asked if he could stop his demonstration. After examining the corner of the table he announced he wanted to deliver a different proof that was shorter and more elegant. Another story concerns an examiner, Tissot, who felt that Poincaré was going to do too well, so he suspended the examination for three-quarters of an hour in order to come up with a harder problem, which he found in an advanced geometry book. Poincaré listened and then drew the figure. He stood and stared at the ground for quite a long time and then started with a solution using advanced methods. Tissot announced that he wanted a more elementary solution. Poincaré then left the blackboard and delivered a trigonometric solution right at Monsieur Tissot's desk, in his face as it were, which prompted Tissot to say, "I desire that you not leave elementary geometry." Poincaré complied and the examiner was happy(and by this time undoubtedly happy to be done). He gave Poincaré the highest grade.

At the Polytechnique, Poincaré continued to sit in lectures without taking notes. One preferred method of working was to march up and down the halls of the school deep in thought. Sometimes in the off hours he would join friends who had also come from Nancy. They would all join arms and walk and talk, but Poincaré would also often sit without speaking or moving, lost in his own thoughts as his friends heatedly discussed the events of that "epoch of crisis".

The Polytechnique's requirements included physical exercise, gymnastics, drill with arms, and drawing. Poincaré did awfully at all of these. His friends were particularly amused by his attempts at drawig and put together a public exhibit of his drawings, with labels in Greek such as "This is a borse." Sometimes they were wrong. "Young

men are without pity like children," Darboux says. In a much-repeated story, Darboux says Poincaré's drawing(*dessin*)was so bad that he scored a zero in his examination for the École Polytechnique. Normally this would stop the entry of a candidate. However, the zero was changed to a 1, which is what appears on the official paper that survives. This story is not quite correct. He received a 1 in *lavis*—wash or tint. Under the category of *dessin*, or drawing, Poincaré received a 12 out of 20. Despite his difficulties in these areas his margin of victory on the total entrance exam was about 7% over his nearest competition. His clumsiness in drawing figures later did cause him fall out of first place when he graduated, because he fouled up the drawing of a figure during the geometry exam and annoyed the examiner. We don't hear of Poincaré being overly concerned with the exams. He was governed by an internal standard and internally regulated interests.

Upon finishing at the École Polytechnique, Poincaré enrolled at the school of mines and gave every appearance of becoming a mining engineer. Many mathematicians and theoretical physicists head off to college with the intention of studying engineering, which seems solid and practical to parents. We have seen Cantor and Koebe, for example, at least momentarily on an engineering track, and the physicists Dirac and Feynman come to mind. However, most are on a career path in mathematics or physics when they leave college. Here, however, we have Poincaré completing school, Passing a postgraduate engineering course, and then working as a mining engineer. He was happy and self-confident, and he was expected to be productive. He performed the duties of his job well. From 1875 to early 1879 Poincaré was mostly in Paris. He made trips, including one to Hungary, to investigate mining techniques. In early 1879 he was sent to Vesoul, seventy-five miles south of Nancy, as a practicing mining engineer. When there was an explosion in a coal mine that killed sixteen people, Poincaré promptly descended the mine shaft to inspect. Darboux says that he showed considerable "*sang-froid*" in his love of duty.

During this period as an engineering student and engineer Poincaré's interest and capabilities in mathematics grew, and he followed a parallel track at the University of Paris, preparing for a Ph. D. degree. In 1878 he published his first original paper, which was on differential equations, and submitted a more substantial thesis, also on differential equations, for the degree. (Poincaré did publish a paper in 1874 while a student at the École Polytechnique, on geometry, but it was confined to giving new proofs of already established results.) Darboux was assigned to read the thesis and was impressed—there were enough original ideas for many theses. He found many points, how-



ever, that needed clarification or revision. Poincaré would never be very concerned with details unless he felt they were important. Darboux reports that Poincaré was accommodating about making corrections that were asked of him. He later told Darboux that his mind had already moved on to other things. The thesis was published in 1879, and Poincaré received his doctorate in August of that year, while he was working in Vesoul.

Poincaré received an appointment to teach at the University of Cean in December 1879, though some duties as an engineer lingered on. He was a promising young mathematician who had done some original work on differential equations. But this first work was like the odd rumblings that might precede the eruption of a volcano, bucause once started, only a word like "eruption" can capture the work that came in the next three or four years.

At the beginning of May 1880, Poincaré read a paper by Fuchs about a spectific kind of differential equation. On May 28 the Academy of Sciences in Paris received from Poincaré a paper that contains many of the key ideas of the theory of automorphic functions, as an entry for its 1880 prize competition. John Stillwell writes, "Poincaré begins in the wilds of differential equations. After a long hack through this jungle, he finally emerges into the light of geometry." (In this paper he drew a figure that was incorrect, noticed it, and told the reader to ignore the part of it that was wrong. He did not recopy the page and draw the figure correctly.) Poincaré himself supplied a piece of



Jules Henri Poincaré (courtesy of the Archives of the Mathematische Forschungsinstitut Oberwolfach)

what his life was like at this time and what his path was to full production as a mathematician. He wrote in *Science* and *Method*, published in 1908:

For fifteen days I strove to prove that there could not be any functions like those I have since called Fuchsian functions. I was then very ignorant; every day I seated myself at my work table, stayed an hour or two, tried a great number of combinations and reached no results. One evening, contrary to my custom, I drank black coffee and could not sleep. Ideas rose in crowds; I felt them collide until pairs interlocked, so to speak, making a stable combination. By the next morning I had established the existence of a class of Fuchsian functions... I had only to write out the results, which took but a few

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

hours.

The I wanted to represent these functions... I asked myself what properties these series must have if they existed, and I succeeded without difficulty...

Just at this time I left Caen, where I was then living, to go on a geologic excursion under the auspices of the school of mines. The changes of travel made me forget my mathematical work. Having reached Coutances, we entered an omnibus to go some place or other. At the moment when I put my foot on the step the idea came to me, without anything in my former thoughts seeming to have paved the way for it, that the transformations I had used to define the Fuchsian functions were identical with those of non-euclidean geometry. I did not verify the idea; I should not have had time, as, upon taking my seat in the omnibus, I went on with a conversation already commenced, but I felt a perfect certainty. On my return to Caen, for conscience' sake I verified the result at my leisure.

Then I turned my attention to the study of some arithmetical questions apparently without much success and without a suspicion of any connection with my preceding researches. Disgusted with my failure, I went to spend a few days at the seaside, and thought of something else. One morning, walking on the bluff, the idea came to me, with just the same characteristics of brevity, suddenness and immediate certainty, that the arithmetic transformation of indeterminate ternary quadratic forms were identical with those of noneuclidean geometry.

Returned to Caen, I meditated on this result and deduced the consequences... [but ran into a problem].

Thereupon I left for Mont-Valérien, where I was to go through my military service; so I was very differently occupied. One day, going along the street, the solution of the difficulty which had stopped me suddenly appeared to me. I did not try to go deep into it immediately, and only after my service did I again take up the question. I had all the elements and had only to arrange them and put them together. So I wrote out my final memoir at a single stroke and without difficulty.

This is a striking passage. Poincaré seems to have an official workday for mathematics of one or two hours during this period. There is also remarkable evidence of self-confidence. When he sees the answer to major difficulties he has other things to do, so he does not verify his insight immediately. He has seen the answer.

Poincaré's creation of the theory of automorphic functions secured his reputation as a mathematician (though it did not win the 1880 prize competition). He published two brief accounts of his results in *Comptes Rendus* in 1881 and was called to a position at



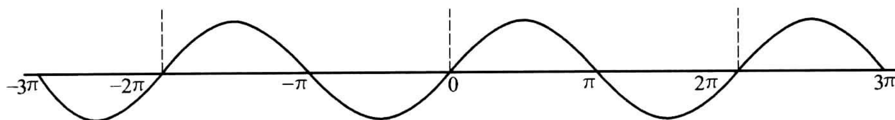


the University of Paris in the same year. He published a full account of his results in a series of four papers in *Acta Mathematica* in 1882-84. Poincaré's theory of automorphic functions and the discrete groups that accompany and generate them are the ground of Hilbert's twenty-second problem.

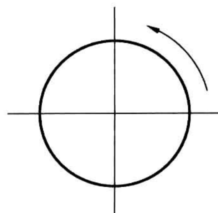
At the beginning of the nineteenth century there was a relatively small roster of functions in the world of mathematics. Algebra supplied the bulk of them, but there were also trigonometric functions like sine and cosine, the exponential function, and special functions created by calculus and/or physics—all functions of real numbers. As the nineteenth century progressed and the real numbers were understood and the imaginary numbers accepted, there was a population explosion in functions. Function theory is an attempt to order, understand, and contribute to this explosion.

What is a function? At first the idea was intuitive. A function took one number as an input and gave back another number as an output. One could draw the graph of a function. Usually a function was identified with a specific rule by which it could be calculated. A simple example is $f(x) = 2x + 3$. As x changes, the result of multiplying x by 2 and adding 3 changes. Sine(x) has a more complicated rule. In its modern, set-theoretical formulation a function is a map (it could be totally arbitrary) from a set that is the domain to a set that is the range. It doesn't have to have any smoothness or even have a method of calculation. The only constraint is that it take each member of the domain to a unique element of the range.

We can illustrate a great deal by examining what happened to the sine function.



Its graph repeats itself. We could build the whole graph by repeating the section from 0 to 2π. One can come up with the sine function in many ways. However, the historical source and the most natural way to get the sine function is to look at a circle with radius 1:



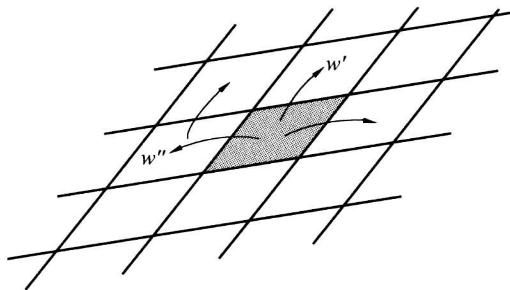
We place the circle in a coordinate system and then walk around the circle counter-

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

clockwise. If we go 2π around the circle we are back to where we started. (We could use degrees, but mathematicians usually use the distance-walked setup.)

The sine function is naturally understood by the picture of the circle and from its first appearance was understood in this way. If we had historically come up with the sine functiion in some other way, it would have been a surprising moment of clarity when someone drew this picture and pointed out the connection. In the first half of the nineteenth century something like this happened with what are called elliptic functions—not when the acquired the name "elliptic", but when they were later connected to the torus. They were discovered when mathematicians tried to calculate the distance around an ellipse, or a portion of the curve of an ellipse. Planets orbit the sun in elliptical paths, so this was a natural thing to try to do. These calculations led to an integral① that could not be solved using any function that was available off the shelf. The name "elliptic" stuck for the function they discovered, but eventually mathematicians were investigating elliptic functions that had very little to do with an ellipse.

Elliptic functions are analytic functions on the complex plane that repeat, like the sine function, but in two separate directions, which we can call w' and w''. Just as $\text{Sine}(x) = \text{Sine}(x + 2n\pi)$ for any integer n, an elliptic function $f(z)$ follows the law: $f(x) = f(z + mw' + nw'')$ for any integers m and n. We can't draw graphs of such a function (domain together with range) because the domain has two dimensions, as does the range, making a total of four. We can draw the domain, though, and mark the pattern of repetition:

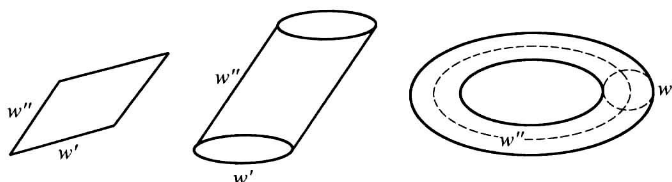


The shaded area is repeated all over the complex plane. A region that can be copied and used to tile the whole plane is called a "fundamental region". Just as the sine function "lives" on a circle and we walk around the circle to get the whole function, a similar thing can be done here. We can cut out one fundamental region—a parallelo-

① This integral was of the form: $\int \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$.



gram—and glue the edges to each other where the function starts to repeat, first getting a cylinder for w'' glued to itself and then a torus(donut) when w' is glued:



We go around one dotted circle(through the hole of the donut)and the function repeats with each revolution. We go around the other dotted circle(walking the top surface of the donut)and the function repeats. Riemann noticed this fact about elliptic functions in his epochal thesis of 1851, which opened up the full range of complex function theory for the first time.

One of the central, remarkable facts about complex analysis is that if we have a little piece of an analytic function we can continue it, building a function uniquely determined in all its particulars by the scrap we start with. As we continue, the function(and its domain of definition)grows to its own pattern, as its own thing, and we are inclined to think at first that we are wandering in the complex plane, because that is where analytic functions originally lived. We wander on our creation in what we, by local landmarks, think to be a circle and assume we are back to where we started. However, the function has a new value. This is curious and difficult, because functions must give only one value for a given input. Riemann's solution was to say we don't really return to where we would calculate we are, but are instead wandering on a more complex Riemannian surface that is the natural home of the function we are investigating. Sometimes it has interconnected leaves, a bit like a screw that we are winding up or down. One of the main tasks of the theory of functions from this time on was to determine the shape and nature of the surface on which a given analytic function of a complex variable lives. The elliptic functions, which were the clue to all this, repeat so reliably that with them all we need is a torus.

Poincaré was blissfully unaware of Riemann's ideas about surfaces when he began to work on automorphic functions almost thirty years later. In the Words of Jean Dieudonné:

Poincaré's ignorance of the mathematical literature, when he started his researches, is almost unbelievable. He hardly knew anything on the subject beyond Hermite's work on the modular functions: he certainly had never read Riemann, and by his own account had not even heard of the "Dirichlet principle", which he was to use in such

imaginative fashion a few years later.

Freudenthal gives a more detailed account of Poincaré's "*ignorance merveilleuse*" in his piece in the volume celebrating Poincaré's centennial. (This volume also contains praises of Poincaré's patriotism, a photograph of his family home, a reproduction of early exam results, and a facsimile of his handwriting.) Poincaré himself says laconically, "I was then very ignorant."

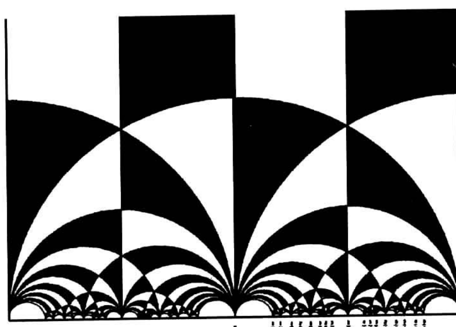
The modular functions Poincaré had encountered in Hermite's work "repeat" in a more complex way than elliptic function under a group of nonlinear transformations of the form $z \rightarrow az + b/cz + d$ where a, b, c and d are integers and $ad - bc = 1$. Fuchs's 1880 paper precipitated Poincaré's work on automorphic functions by leading him to investigate questions about a type of differential equation and the functions that come from it. Assuming they existed, these functions would remain invariant under discrete groups of transformations of the form $az + b/cz + d$ with a, b, c and d allowed to be real numbers. In a discrete group no transformation (be it a member of the group or not) can be arbitrarily well approximated by other members of the group. Just as in the case of elliptic functions, where there is a fundamental region that, repeated, covers a domain, the groups that Poincaré was thinking about in his work on automorphic functions take a more complicated polygonal region (a fundamental region) and cover a domain, which for automorphic functions is the upper half of the complex plane.

Poincaré came at this from the direction of equations and algebraic transformations, so he did not have the clear drawings we have today, in a way analogous to (but far more complicated than) not knowing how the sine function relates to a circle. This was early 1880s work. One may be led to think of crystallographic groups, but Hilbert referred to Poincaré's work in setting up the eighteenth problem.

As we saw in the passage quoted, after drinking too much coffee Poincaré saw how to construct functions that remain invariant under a group of these transformations. It was in the air that a discrete group of this type has a fundamental region that is reproduced in a distorted way across the upper half-plane. As he put his foot on the step to the omnibus in Coutances, it came to him that these algebraic transformations were of the same form as transformations that appear in non-euclidean geometry. Here is a drawing of this for the modular group:

He realized that this picture (and any more complicated picture of a Fuchsian group's division of the upper half-plane) was a distortion of a whole non-euclidean plane and the fundamental regions were exact copies of each other. It is particularly amazing that Poincaré was able to "see" this coming from the direction of the equations.





Tessellation of the modular group. A dark and a light region together make up a fundamental region.

With the insight that non-euclidean geometry provided, Poincaré proved a wide range of results. He showed that if one starts with a polygon that has appropriate geometrical properties, one can then find a Fuchsian group that generates it. He discovered a classification system and explored some examples in detail. He observed that the edges of a fundamental region are naturally connected to each other by the action of the generating group. Therefore he imagined cutting out the fundamental region and gluing the edges together, just as the edges of the parallelogrms are glued together in the case of elliptic functions to get a torus. He quickly observed that the resulting two-dimensional surfaces were topologically complex and could not exist embedded (without distortion) in normal threedimensional space. These more complicated surfaces are the natural home for automorphic functions, just as elliptic functions live on a torus. Poinicaré proved that if one starts with a compact fundamental region one obtains a surface that is characterized topologically by its genus—the number of holes it has. He had rediscovered Riemannian surfaces. By 1881 Poincaré was corresponding with Klein, who had been working on this problem also, and Klein informed him of the existing literature, so he didn't have to reproduce all of Riemann's work. There was a rivalry between Poincaré and Klein in this period, felt more strongly by Klein.

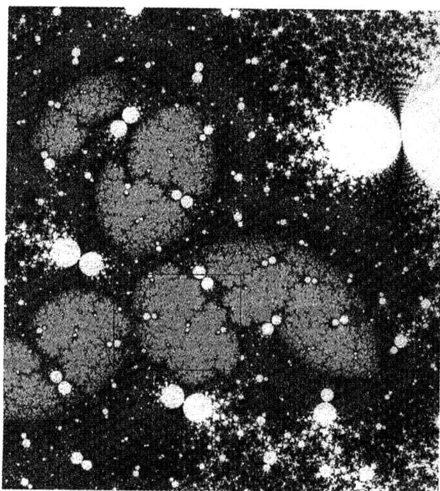
Poincaré published his work on discrete groups in 1882 in *Acta Mathematica* in a substantial setting of the stage. In his second *Acta paper*, he showed how to construct an automorphic function for any discrete group. This paper is a virtuoso performance of classical analysis. Then came a third *Acta* paper exploring a generalization of Fuchsian groups that Poincaré named in honor of Klein. (When Poincaré honored Fuchs by calling automorphic functions Fuchsian, Fuchs hadn't worked on them. Now, Klein had not yet worked on Kleinian groups.) A Kleinian group is like a Fuchsian gruop except that a, b, c and d are now allowed to be complex numbers instead of real numbers. This has

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

large consequences, and Stillwell put it, writing in 1985:

A Kleinian group action on the plane turns out to be just the shadow of a discontinuous group of three-dimensional non-euclidean motions... the difference betwwen Fuchsian and Kleinian can be understood as the difference betwwen two and three dimensions. This is of course a vast difference, and the theory of Kleinian groups after Poincaré had to wait for further progress in three-dimensional geometry and topology, much of it coming only recently.

Here is part of a picture of the limit set of a Kleinian group drawn with the aid of a computer. (The limit set is the set of points that are approached by repeated actions of the group on the points in the domain.)



Limit Set of Kleinian Group (courtesy of Peter Liepa)

Poincaré's fourth *Acta* paper on automorphic function is the most open-ended, suggestive, and least rigorous. In a paper like this one, and there were many others, one can get an inkling of how deeply Poincaré was looking into the future of mathematics, as if he were building a ship in a bottle through a wormhole to the future. Part of this paper is about uniformization by automorphic functions, the subject of Hilbert's twenty-second problem.

During the same time period as the papers on Fuchsian functions, a long paper (also in four installments), "Memoir on the curves defined by a differential equation," was conceived, written, and guided to publication. Many of the ideas were already present in Poincaré's thesis. Darboux informs us that one of the reasons Poincaré was able to write so fast was that he never revised. (Hilbert, on his visit to Paris as a young man,



expressed reservations about Poincaré's work partly for this reason.) Dieudonné says:

The most extraordinary production of Poincaré, also dating from his prodigious period of creativity (1880 ~ 1883) (reminding us of Gauss's *Tagebuch* of 1797 ~ 1801), is the qualitative theory of differential equations. It is one of the few examples of a mathematical theory that sprang apparently from nowhere and that almost immediately reached perfection in the hands of its creator. Everything was new in the first two of the four big papers that Poincaré published on the subject between 1880 and 1886.

A differential equation relates the rate of change of something (e.g., physical position) to something else (e.g., force applied). The most famous of all differential equations in Newton's $F = ma$, force equals mass times acceleration. The force in this equation is specified in some way, the mass is the mass of the body being pushed, and acceleration is the rate at which the velocity of the body being pushed is changing. In turn, veloctiy is the rate at which the position of the particle is changing. The goal is to move from this equation, which involves acceleration and force in the moment, to a description of the movement of the body as time unfolds. We can write differential euqations for many things—fluid folw, the orbits of the planets, the propagation of electromagnetism, the flow of heat, or even the growth of a population of fleas. The equations for the sun, earth, and moon cannot be solved exactly, and those for more complex situations are even more difficult. The central mathematical tool of science, differential equations have been studied and studied hard. They were clearly the center of Poincaré's interest in mathematics. Even when he does epoch-making work in geometry or topology, we find he was usually led there by a differential equation. Poincaré's first paper was on differential equations, his last paper was on differential equations, and he wrote at least one paper on differential equations just about every year of his mathematical life (as Dieudonné observed). His work on celestial mechanics is about differential equations. Poincaré's universality as a mathematician did not lead him to differential equations. Rather, it could be argued, his interest in differential equations flowered into universality.

The problem with differential equations is that most of them are too difficult to solve. When one attempts to calculate the length of a curve, or the area bounded by a curve, at first everything works wonderfully (as it does in a first-year calculus class). Investigating a circle magically leads to the trigonometric functions (which were already available), and they in turn supply the number π. However, as we have seen, the system already breaks down when one attempts the length of a part of an ellipse.

In 1880 a few elementary differential equations were solvable exactly, and there

were some existence theorems that asserted the existence of solutions within a region but did not specify them. There were also procedures, usually of questionable validity, that allowed one to grind out an approximate solution in a given area using a series expansion of some kind. Some of these were clever and worked well in practice. Today one can buy software that will "solve" differential equations on a personal computer. These programs often make use of these old methods.

Poincaré set out to see what he could figure out about the global solutions of a general type of nonlinear differential equation, specifically, in Poincaré's terms, $dx/X = dy/Y$, where X and Y are polynomials and only real numbers are allowed. He started out by looking at the "critical" points of the equation, that is, when X and Y were both zero and the equation lost its meaning (no matter how one wrote it) because it involved dividing by zero. Cauchy and Charles Auguste Albert Briot (1817—1882) had already looked at these points and classified them as "nodes", "saddles", "spiral points", and "centers". However, Poincaré addressed the problem in wholly original terms. In order to get a picture of curves that ran off to infinity, he projected the plane onto a half-sphere sitting above it, then flattened the sphere so that everything could be drawn on a disc. Here we have one of those jumps so striking in the work of Poincaré. Suddenly the space one is working in is non-euclidean, and geometry and topology jump into a problem on differential equations. One is discovering how to fit things in and have the ends match up.

Poincaré generalizes the problem; the habital for the curves being sought is a surface determined by a polynomial. Early topological invariants like genus and Betti numbers come into play. One way of viewing the problem when phrased in this way, and how geometry and topology get into it, is by viewing the vector field that is defined by these equations as hair growing on the surface that has to be combed and to lie flat. The vectors come out of every point on the surface (as the hair on one's head does not). The nodes and spirals, etc., represent different kinds of cowlicks. They are curcial in the behavior of any curve (or solution of the differential equation). In general, it is impossible to comb a hairy surface without getting some cowlicks. An exception is a torus.

The hair is combed round and round in the same direction.

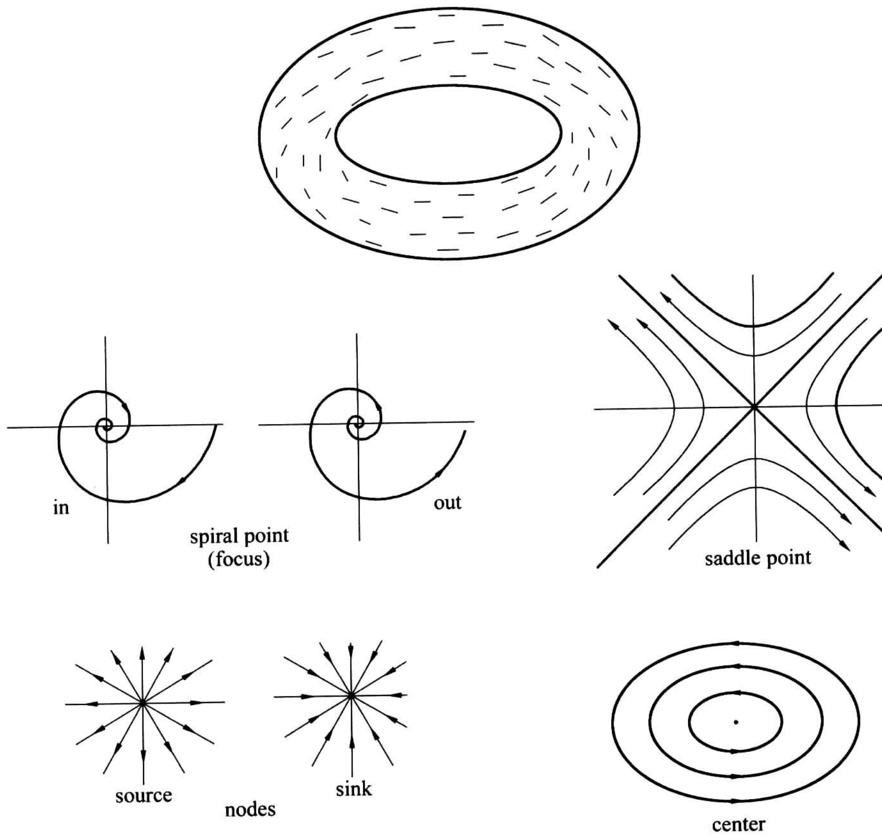
Types of critical points would include:

(You might get a saddle if you aimed two streams of water at each other head-on.)

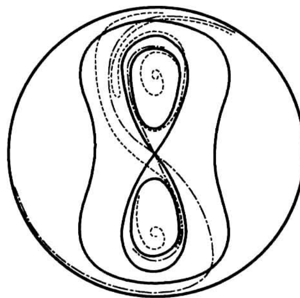
Poincaré introduces the genus, or number of holes in the surface, and then proves a remarkable result:

$$\text{Genus} = (\text{Saddles} - \text{Spiral Points} - \text{Nodes} + 2)/2$$





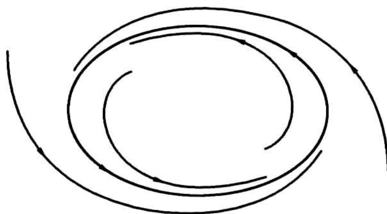
He showed that centers, though common with very simple equations, occur only in exceptional cases when the equations are more complex. Something like what is now known as the ergodic hypothesis makes an appearance, though Poincaré was unable to prove it. If one looks at some of these illustrations Poincaré included after having followed the popularization of chaos theory and the many diagrams that have a similar feel, one is haunted by how deep Poincaré's vision into the future was.



A qualitative picture of the solution to a differential equation. From Ouevers de Henri Poincaré (vol.2, p.71)

In the last section of this paper, published in 1886, Poincaré looked at the problem in higher dimensions. He had already started looking at the three-body problem and the problems of celestial mechanics.

Poincaré was working on the qualitative theory of differential equations. The second half of Hilbert's sixteenth problem is a question in this area. The nodes(sources and sinks), spiral points, and saddles are all defined in terms of how solutions approach a singularity—a singularity, in this context, is where the function is not defined. Farther out in the region of a singularity one might find a limit cycle. A limit cycle is a periodic solution—a closed orbit—that has an"annulus-like" neighborhood in which there are no other periodic solutions. One example could be drawn:

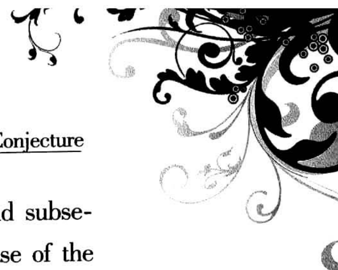


These are a bit like eddies. This particular limit cycle attracts nearby solutions, which approach it asymptotically. Hilbert simply want to know how many limit cycles are possible for a given equation.(This is one of the unsolved problems.)

In 1923 the French mathematician Henri Dulac"proved" the number of limit cycles for these equations was finite. In 1957 Petrovsky and E. M. Landis(1921—1997)presented an estimate of how large this number was. The"solution"to the sixteenth problem Petrovsky and Landis published in 1957 was heralded. However, a gap in their argument was soon discoverd, resulting in a retraction some years later. The correctness of the bound or limit Petrovsky and Landis found remained in a kind of limbo. In 1980 Shi Songling published a counterexample. Then in 1981 Yulij S. Ilyashenko found a gap in Dulac's proof, so that not only was the number of possible limit cycles not known, but it wasn't even certain that the number was finite. Finally, Ilyashenko(1991)and Jean Ecalle(1992)independently, and using different methods, gave rigorous proofs for Dulac's finiteness result, employing radically more sophisticated methods than Dulac had at his service.

This question is seen as important today. Smale included it on his list of problems in 1998, which implies some optimism that it might be solved. If we could better understand the second half of the sixteenth problem, it would be a step in the direction of better understanding the weather or other complex systems that can fall into repeating cycles. I asked Ilyashenko how things stood in late 2000. He was optimistic but didn't





expect a solution soon. He said that a paper by S. Yakovenko and himself and subsequent work of V. Yu. Koloshin provide a "Hilbert type" bound to a special case of the sixteenth problem. This is significant because it is the first time any bound of this type has been found.

In 1885 King Oscar of Swden proposed an international competition, with the goal of finding a convergent series solution to the n-body problem. With a good solution one might prove the stability of the solar system. This was a perfect opportunity for Poincaré. Mathematicians including Pierre-Simon de Laplace(1749—1827), Lagrange, Siméon-Denis Poisson(1781—1840), and Dirichlet all claimed to have proved the solar system stable. Dirichlet died before he could write down his proof, and the proofs that were written down were not convincing. These efforts mostly boiled down to approximate calculations of the movements of the planets based on infinite series that might not converge and trying to extend that indefinitely into the future. These expansions, though often useful for practical problems, invariably contained, sooner or later, terms with "small denominators" that could disturb whatever orderly pattern seemed to be unfolding in the lower-order approximations. However, these methods had been for years producing reliable calculations of where the planets should be.

Poincaré asked a great question: Why do these defective methods often work as well as they do? His answer was that many of the series involved were asymptotic. He showed that these series generally do not converge to any one answer, but they diverge so slowly that if we use them as approximations, the error, though it doesn't disappear in absolute terms, grows ever smaller on a percentage basis. Further, the percentage of error often becomes small after a small number of calculations. Poincaré—and independently Jan Stieltjes(1856 ~ 1894)—inaugurated the general theory of asymptotic series in 1886 papers.

Using a brilliant array of classical techniques with his clarifying theory of asymptotic series, as well as his new qualitative picture, Poincaré attacked the threebody and n-body problems. He proved rigorously that past "proofs" of stability in celestial mechanics were invalid. The series did not uniformly converge. Yet they were good mathematics. The three-body problem is complex, has many versions of exactly how it is stated, and, at best, starts from an approximation(point masses in Newton's equation). This qualification of the word "proof" is essential to understanding what has happened in the area. Poincaré's writings began a new era in celestial mechanics in which the statements of theorems and their proofs could be judged by more rigorous standards. In his first paper in this area in 1883, predating the prize, Poincaré had shown that there were an infinite

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$

number of solutions to a restricted three-body problem of a type G. W. Hill(1838—1914)discovered in 1878. Neither Hill nor Poincaré actually proved the convergence of their answers. Poincaré returned to the restricted three-body problem and proved this time that if the mass of the second body was small enough, there were periodic solutions. Each member of the growing array of special solutions to the restricted three-body problem(including Poincaré's) was not proved stable if jostled in an tiny way.

Though not a solution to the question about the stability of the solar system, Poincaré's progress understanding the equations involved was substantial and he won King Oscar's Prize in 1889. Poincaré proved the existence of"homoclinic"points, points of stability that can be approached from more than one distinct direction and left in more than one direction. Disturbed in small ways, the system becomes wildly unstable—in modern terms, chaotic. The picture of the path of solution becomes a homoclinic tangle. Poincaré said he wouldn't try to draw the picture.

It has recently been found by scholars that the 273-page paper published as the prizewinner wasn't, in fact, exactly the one turned in. After the journal was already in print, errors were discovered. Poincaré went back and found that the error was deeper than he had initially thought. He ended up paying to have the issue reprinted, which exceeded the prize money he won. But the end result was good—the paer finally published was correct and Poincaré had discovered what we now call chaos theory in the course of his own somewhat chaotic process.

By the mid-1890s Poincaré had outrun the mathematical vocabulary of his time. For example, in attempting to formulate the question of stability, he was driven to talk about what we would now call measure theory. He ended up using the vocabulary of probability, which was scarcely more clearly grounded than the as-yet-nonexistent measure theory. Poincaré did not neglect calculation in his assault on celestial mechanics. His treatment of the classical methods of approximation in celestial mechanics is still being recommended, for example by V. I. Arnold in *Mathematical Methods of Classical Mechanics*(Russian 1974, English 1978). Poincaré wrote close to 100 papers on classical computational topics in the theory of the solar system.

Poincaré worked on many other physical problems, and he continued to be unafraid of difficult nonlinear problems. One of his most famous results was the stability of a rotating pear-shaped body of fluid held together only by gravity—this suggests the possibility of a pear-shaped star. He wrote and thought about the partial differential equations of physics—heat, fluid flow, elasticity, etc., which were at a very simple level of development in his time. Physicists today are likely to encounter methods associated with





Poincaré in these areas. Writing about chaos theory, L. Galgani says:

Obviously, no revolution is actually fully a revolution. Indeed, in a sense almost everything can already be found in Poincaré; the homoclinic phenomenon leading to chaos is explicitly described; the possibility of the existence of invariant tori with good frequencies satisfying suitable diophantine conditions is explicitly considered and it is shown that one cannot exclude it...

In this connection, it is of interest to remark that possibly the present state of awareness can be understood if one takes into account the role... [of high speed numerical computers]. Indeed the speculations of Poincaré on homoclinic points were probably too advanced, even for the community of pure mathematicians, and only after the exhibition of their effects by Hénon[1964]did people become acquainted with them.

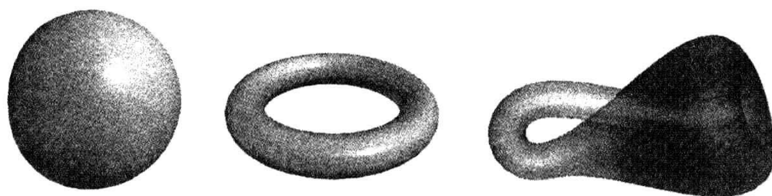
Topology particularly begged for development. Poincaré said, "Every problem I had attacked led me to *Analysis situs*." The idea and words for an *analysis situs or geometria situs* go back at least as far as Leibniz, who wanted an inherently geometrical analysis of space that was as powerful as Descartes's analytic geometry, which used algebra. Despite Leibniz's early call for order, as of 1894 there was no organized discipline of *analysis situs*(or topology), just isolated and mostly intuitively derived special results. Riemann had introduced fundamental topological concepts, such as the degree of connectedness(how many cuts are required to cut a surface into separate, simple pieces)and the idea of a manifold(a surface that is locally similar to euclidean space). Enrico Betti(1823—1892)generalized the idea of connectivity to higher dimensions, but without a general foundation. Johann Listing (1806—1882), Augustus Ferdinand Möbius(1798—1868), and Klein discoverd odd, one-sided surfaces. Poincaré had spent a major part of his career already dealing with topological questions and had had enough of this uncertainty. In 1895 he created algebraic topology—though it would not be commonly called topology until the 1920s and became heavily algebraic only with time.

As Poincaré himself described the subject:

Imagine any sort of model and a copy of it done by an awkward artist: the proportions are altered, lines drawn by a trembling hand are subject to excessive deviation and go off in unexpected directions. From the point of view of metric or even projective geometry these figures are not equivalent but they appear as such from the point of view of geometry of position[that is, topology].

Roughly, topology attempts to answer the question of which shapes are the same if we allow stretching but no tearing. Was his youthful effort a drawing of a horse? For example, the sphere, the torus, and the Klein bottle(these are surfaces, not solid ob-

jects):



How are they different? The torus has a hole, while the sphere doesn't. That distinguishes them. Both the sphere and the torus have an inside and an outside to their surfaces. The Klein bottle doesn't. These ideas were known. The question facing Poincaré was how these insights could be made precise and generalized. He came up with two separate approaches—what we would today call homology theory and the first homotopy group(also known as the fundamental group). Dieudonné wrote: "It has been rightly said that until the discovey of the higher homotopy groups in 1933, the development of algebraic topology was entirely based on Poincaré's ideas and techniques."

Poincaré had married Louise Poulain the year he was called to Paris, 1881. She was the great granddaughter of Étienne Geoffroy-Saint-Hilaire, who was the *antagoniste* of Cuvier, a naturalist famous for pathbreaking work on fossils and comparative anatomy. Despite their scientific disagreements, Saint-Hilaire was responsible for Cuvier's obtaining a position at the national museum. His descendants continued the association with the museum, and Darboux reports that Poincaré's mother-in-law actually lived at the museum for a long time. Poincaré and his wife had three girls and a boy. Much of Poincaré's extended family was in Paris either permanently or visiting. He saw much of his sister and her family, the Boutrouxs. Photographs of family gatherings or vacations are full of children. One imagines that the type of life Poincaré lived as a child continued into his adulthood. His fame and genius coupled with his mild, distracted disposition would have allowed him to be the focus as much as he wanted to be.

During the 1890s and the first decade of the 1900s Poincaré became increasingly lionized. He traveled widely and was invariably a featured speaker at any conference he attended, though Hilbert upstaged him a bit with his list of problems at the 1900 Paris Congress. He traveled to the U.S. more than once. Darboux says he was particularly excited by a trip to St. Louis, Missouri, in 1903. Toward the end of his life his opinion was sought on every conceivable subject, from politics to the effect of comets on the weather. In the words of Dieudonné, offering popular explanations of mathematics and science was "a chore for which he does not seem to have shown great reluctance." He received honors and titles and served on committees for railroads, the post and telegraph, the national observatory, and educational institutions. Three times, in 1899, 1909, and



1910, he was president of the *Bureau des Longitudes*. He was a member of at least twenty-one foreign academies and many foreign societies, received at least nine honorary doctorates, and was active in the cause of international bibliography. Poincaré never actually ceased to be a mining engineer, but remained under the control of the department of public works; he was just on assignment at the universities. In fact, in 1893 he became chief engineer of mines and in 1910 inspector general. Of course, he won many prizes.

Poincaré wrote three popular books on mathematics and science: *Science and Hypothesis* (1902), *The Value of Science* (1905), and *Science and Method* (1908), from which we quoted his account of his discover of Fuchsian functions. The books mark a high point in writing about science and help explain his appointment to the French Academy as an arbiter of the language. Scholars in the history of science sometimes refer to these books as primary sources, in particulr for the comments they contain on time, the aether, geometry, and the ideas behind chaos theory. According to Abraham Pais, Einstein had read parts of the first in the years preceding 1905 and the birth of the theory of special relativity.

There has been some debate whether Poincaré should receive equal credit with Einstein for the creation of special relativity, based on a paper he published in 1905 as well as for his clear, early statement of some of the basic principles in the popular books and other places. It is true that no one was more conversant with the mathematical theory of electromagnetism than Poincaré. The difficulties that require the special theory were known to him. The mathematics of the solution must have been child's play. But what did it mean physically? Poincaré had all the pieces, and in 1905 he did publish a theory of relativity that was similar to Einstein's. However, Pais makes a convincing argument that Poincaré didn't, in 1905 or ever, fully understand the way special relativity flows from two simple hypotheses, that light always travels at a constant speed and that the laws of physics remain the same in any inertial frame of reference. As late as 1909, in one of his lectures in Göttingen, Poincaré was still referring to a third hypothesis—lenght contraction, which follows from the other two. It therefore seems an exaggeraton to give Poincaré equal credit with Einstein for special relativity.

Poincaré continuously wrote papers on ideas that occurred to him—on number theory, algebra, and myriad special topics from anslysis. His work on abelian functions and algebraic geometry occupies almost as much space in his complete works as automorphic functions. However, he was not the clear leading figure in algebraic geometry in this period. (Since this is Poincaré, in another fifty years we might find out this is his deepest work.) His most important paper on number theory (1901), Dieudonné says, "was the

first paper on what we now call 'algebraic geometry over the field of rationals.'" This is ground zero for "arithmetic algebraic geometry." Mordell's 1921 theorem and Siegel's 1929 results on diophantine equations start here. The mathematician Isabella G. Bashmakova notes: "It seems that Poincaré was completely unaware of the work of his predecessors bearing on the arithmetic of algebraic curves. He knew of Diophantus' procedures."

The last of Poincaré's series of papers on *analysis situs*, in 1904, marks an end to the period of expansive creation that began on the omnibus in 1880. However, nothing is so cut and dried with Poincaré. The key in much of his last work is that by this time Poincaré and others had made progress toward creating techniques and vocabularies that allowed questions he might have groped after earlier to be addressed more precisely.

Having invented algebraic topology, Poincaré could now return to the topic of uniformization of surfaces defined in terms of algebraic and analytical equations. His original 1884 attempt at a proof hadn't been complete. Hilbert (though he seems to think Poincaré had earlier been more successful than he really had) called completing this project his twenty-second problem. Poincaré probably thought of it as unfinished business.

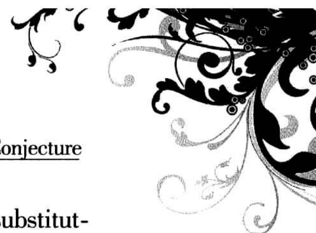
We can now express the problem simply, if we bear in mind that to uniformize is to parameterize a surface. The surfaces to be uniformized might be generated by polynomial equations or in some other way. Hilbert want to prove there exists a parameterization using automorphic functions of any surface defined by "any analytic non-algebraic relations whatever." In particular, he wants to be sure "all regular points of the given analytic field are actually reached and represented." It is easy to parameterize a piece of such a surface, and Poincaré had already done this in the *Acta* papers. He proved the full result in 1907 using a process of exhaustion and a universal covering space (from algebraic topology), and thereby solved "Hilbert's" twenty-second problem at the age of fifty-three. Koebe solved the same problem independently and pursued his solution in its consequences with more thoroughness the same year.

Why is this seemingly technical result useful? To understand what a parameter is, consider the equation of a circle: $x^2 + y^2 = 1$ (We don't include complex numbers in this simple example). It is easy to specify both x and y as functions of a third variable, t (the parameter):

$$x = \frac{2t}{1+t^2}, y = \frac{1-t^2}{1+t^2}$$

One of the central goals of number theory is to find rational solutions to polynomial equations. Once we have the parametric representation above, it is easy to see how to





find a lot of rational solutions to the equation for a circle. Any rational number substituted for t yields a rational solution. Success in the twenty-second problem makes the automorphic functions a complete set of building blocks for parameterizing all of the large class of surfaces Hilbert asked about, despite the fact that they would seem to be a narrow subset of all possible analytic functions. Remember, it had been quite an achievement for Poincaré to prove the automorphic functions even existed. Now they could stand in for the functions using t in our simple example and provide a potential handhold on myriad new problems. Once again, Hilbert was concerned with stocking the tool chest of mathematics.

Poincaré received his first intimation of mortality in 1908 when he was attending the fourth International Congress of Mathematicians in Rome. He suffered from enlargement of the prostate, which prevented his delivery of a talk on the future of mathematics. He was operated on by Italian surgeons and seemed to recover. Tobias Dantzig gives a sense of his presence in this period:

I saw him often between the years 1906 and 1910, when I was a student at the Sorbonne. I recall above all his unusual eyes: myopic, yet luminous and penetrating. Otherwise, my memory is that of a man small in stature, stooped and ill at ease, as it were, in limb and joint. This last impression was accentuated by his manner of writing on the blackboard...

In 1910 and 1911 Poincaré published in two parts what is probably his deepest paper on algebraic geometry. Then his difficulties returned, and in 1911 he wrote to the editor of that because of his age and health he was sending an incomplete work for publication. He had returned to the three-body problem. The work had given him great difficulties, but he thought it important and wanted his ideas published. This is sometimes known as "Poincaré's last theorem". He succeeded in reducing the problem of the existece of periodic solutions of the restricted three-body problem to a topological theorem about areapreserving mappings of an annulus onto itself, subject to certain boundary conditions. If all such maps had fixed points, that is, points that were mapped onto themselves, then there were periodic solutions to the restricted three-body problem. Dieudonné syas this "was probably the first example of an existence proof in analysis based on algebraic topology". In 1913 a young American, George D. Birkhoff(1884—1944), proved the theorem. Many more theorems of this type have followed. It was a new way to prove the existence of and understand periodic solutions, and a direct descendent is the Moser twist theorem in KAM theory. When one looks at the theory of mechanics today one could mistakenly think one had wandered into an algebraic topology



Jules Henri Poincaré(courtesy of the Musée de la Poste)

forum.

Poincaré did not see the proof of his theorem. On July 9, 1912, he had another surgery that seemed to be successfl, but eight days later he died suddenly, probably because of an embolism. Great mourning ensued. Pall Painlevé—a mathematician turned premier/president of France—called him "the living brain of the rational sciences," which, though dramatic, was not far from the truth.

How could anyone be so productive? One answer comes from Poincaré's nephew Pierre Boutroux, in a letter to Magnus Mittag-Leffler:

He thought in the street as he went to the Sorbonne, while he was attending some scientific meeting or while he was taking one of his habitual grand walks after lunch. He thought in his antechamber, or in the hall of meetings at the institute, while he walked with little steps, his physiognomy tense, shaking a bunch of keys. He thought at dinner table at family get-gogethers, even in the sitting rooms, interrupting himself often brusquely in the midst of a conversation, and would leave his interlocutor planted, to follow a thought that crossed his spirit. All of his work of discovery was done mentally... without him, most often, having to check his calculations in writing.

So the answer is: Think rapidly, think all the time, and write down everything only once, when you are going to publish it. Start with miraculous natural talent. One is drawn to compare Poincaré with Hilbert. Poincaré's relatively early death and Hilbert's long life tend to exaggerate their seven-year age difference. Both were universal mathematicians under E. T. Bell's definition. Both became mathematically productive relatively late—in their mid-twenties—and neither rushed his schooling. Each was viewed as the other's only real rival. Yet they were very different.

Hilbert was disciplined in his writing and, though he didn't always enjoy it, was



systematic in his study of the literature, retaining assistants to do much of the slogginig for him. He worked on one subject at a time, to the exclusion of others. Algebra, number theory, geometry, analysis (including analytic number theory), mathematical physics, foundations. Once done with a subject he seldom went back. In contrast, Poincaré wrote rapidly and seldom revised. When starting his career, he was almost unbelievably ignorant of the existing literature. He eventually came to know quite a lot, but less through systematic research than because he had an active mind, and once he glanced at something he understood and remembered it. Not only did Poincaré not work on only one subject at a time, the subjects he worked on were often hybrids.

We can see some of the point of view Poincaré wanted to leave in his paper "*L'avenir des mathématiques*" ("The Future of Mathematics"). This is the talk he didn't get to deliver at the 1908 International Congress in Rome because of the first flare-up of his illness—perhaps an answer to Hilbert's "Mathematical Problems." Its qualitative character stands out. Poincaré generally doesn't pose specific problems, but talks about directions. He discusses the importance of the invention of new words and the importance of words in mathematics. The word "energy" is an example, and he calls it prodigiously fecund, comparing it to a machine that economizes effort. He talks in terms of the elegance of methods of thought and of "the art of giving the same name to different things" when what we say about one thing is equally true of the other. Where a mathematician less sensitive to language might think of dual theories, Poincaré uses the word "analogy," and *analogie* or *analogue* appear again and again in the sixteen-page paper. As he thought of the equations needed for automorphic functions and stepped onto the bus in Coutances, he saw non-euclidean geometry. As he thought about finding the sometimes rational solutions to polynomial equations, he saw the separate rational solutions as part of a picture—an arithmetic algebraic geometry. When he thought of differential equations he saw one picture after another. In his work on dynamics, the three-body problem, and nonlinear differential equations, the specific results, important as they are, are probably not as important as the fact that Poincaré recast the mathematical language in which these problems are phrased. The coordinates of where something is located at a given moment supply the same information that a still photograph would and don't tell us the direction and speed things are going. Poincaré included velocities as coordinates of location in what is now called "phase space," and the still photo changes into a picture of a more complicated space that shows the system as flowing from one state into the next, into the next. A specific solution is a clean line or curve in this picture. The idea of hair and cowlicks, the qualitative theory, makes an appearance. This

change of point of view paid huge dividends. Poincaré didn't originate all of the ideas of doing "mechanics on a differentiable manifold"—a string new of words—but he transformed bits and pieces into a systematic vision that he used to investigate hitherto intractable problems. His ideas are still being explored with great energy.

Whe have returned many times to the controversy over Cantor's set theory. Earlier, I glided over Poincaré's intuitionism and said that whatever reservations he may have had about Hilbert's program and set theory, Poincaré was hardly a partisan of a restricted mathematics. In"The Future of Mathematics"he puts his point of view concisely. He seems to be reactig almost instinctively to the limits of merely mechanical language and purely formal systems. For him a machine has levers and grease and pulleys and will let him ride down into a mine. In"The Future of Mathematics"he admits that certain results arising from Cantor's theory are useful. (After all, as we have seen, the words of set theory are a remarkably efficient machine for saying things about mathematics.) But then he says, "I think for my own account, and I am not alone, that the important point is to not introduce things that one cannot define completely in a finite number of words." This almost seems to jump off the page and conjure Paul Cohen's minimal model, in which "every element can be 'named'." In this modes the continuum hypothesis is true, but Cohen then uses it to construct a second model in which the continuum hypothesis is false—proving it independent. Poincaré would have undoubtedly have seen Cohen's method as a more precise rersult, but in his time, more than fifty years earlier, his intuition told him that these questions were imprecise, and he had some intuition why.

This book celebrates Hilbert's problems. In his 1951 article, "A Half-Century of Mathematics," Hermann Weyl wrote:

David Hilbert... formulated twenty-three unsolved problems which he expected to play an important role in the development of mathematics during the next era. How much better he predicted the future of mathematics than any politician foresaw the gifts of war and terror that the new century was about to lavish upon mankind! We mathematicians have often measured our progress by checking which of Hilbert's questions had been settled in the meantime. It would be tempting to use his list as a guide for a survey like the one attempted here.

This chapter on Poincaré suggests some ways in which Hilbert's problems are incomplete. To cover more of the century, Particularly the second half, we should discuss Poincaré's problems, though he didn't leave them in a tidy list. Here are nine such problems touched on in the previous pages and unsolved at the time of Poincaré's death. One could easily find others.



1. Describe the limit sets of Kleinian groups.

2. Investigate what are now called Teichmüller spaces—first suggested by Poincaré in the fourth *Acta* paper on automorphic functions, 1884.

3. Solve the second half of what came to be called Hilbert's sixteenth problem—and more generally extend Poincaré's qualitative methods in the study of differential equations.

4. Achieve further results on of the problem of small denominators in the n-body problem. (KAM theory did this, 1954 ~ 1963.)

5. What is the significance of homoclinic points and homoclinic tangles in the n-body problem? (Chaos theory addressed this in the second half of twentieth century.)

6. Put Poincaré's work in algebraic topology on a firm foundation-homotopy and homology. Explore the immediate consequences of these ideas. (Accomplished by many, 1900—1940.)

7. Prove or disprove the Poincaré conjecture for dimensions three and greater—a three-dimensional manifold is essentially the same as (homeomorphic to) a three-dimensional sphere if the manifold is closed, smooth, simply connected, and has the same homology. (Solved five or more dimensions, Smale, 1960 ~ 1961; four dimensions, Michael Freedman, 1982; three dimensions, unsolved as of 2000.)

8. Explore the ideas in Poincaré's 1901 paper about algebraic geometry over the field of rational numbers to prove results about diophantine equations. This is now called arithmetic algebraic geometry. (Central to Wiles's 1994 proof of Fermat's last theorem.)

9. Prove Poincaré's last theorem—the theorem Poincaré published without proof just before he died. (George Birkhoff, 1913.) More generally, extend this topological method in the study of dynamics and stability. (Work is ongoing.)

附录Ⅺ 陶哲轩怀念威廉·瑟斯顿①

威廉·瑟斯顿(William Thurston),曾对低维流形和相关结构的理解做出过本质性贡献的著名数学家,在本周二(2012 年 8 月 21 日)去世了,享年 65 岁.

或许瑟斯顿最有名的成就是 Haken 流形的双曲化定理,该定理表明满足一定拓扑条件的任何三维流形都可以赋予一个双曲几何结构(即一个能够使得该流形等距于一个双曲的三维球面 H^3 之商的黎曼度量).这个连接着三维流形的拓扑结构与几何结构的高难度定理引导瑟斯顿给出了有广泛影响力的几何化猜想,这个猜想(至少理论上)对任意三维流形的拓扑结构进行了完整的分类,分为八种标准的几何(现在被称为瑟斯顿标准几何).这个猜想有很多推论,包括瑟斯顿的双曲化定理和(最著名的)庞加莱猜想.事实上,它把庞加莱猜想置于一个概念上更吸引人的一般框架中,这个框架中许多其他的情况已经被证实,瑟斯顿提供了一个指向证实庞加莱猜想的最强有力的证据.直到 2002 ~ 2003 年中,佩雷尔曼(Grigori Perelman)的工作通过发展哈密尔顿(Richard Hamilton)的瑞奇流(Ricci flow)的方法证明了庞加莱猜想和几何化猜想都是正确的.(现在有几个佩雷尔曼关于两个猜想证明的变体;在 Besseires, Besson, Boileau, Msilot 和 Porti 的几何化证明中,瑟斯顿的双曲化定理即是一个关键的部分,它容许绕过佩雷尔曼证明中关键的一步,即亚历山大洛夫空间理论)

瑟斯顿结果中我最喜欢的一个是他关于球面反转的优美方法(光滑地将 $\mathbf{R}^3$ 中的一个 S^2 球面内部翻到外面,翻转过程中不能有折叠或者奇点).球面可以内外翻转这一事实是非常不直观的,经常被称为斯梅尔悖论,因为史提芬·斯梅尔(Stephen Smale)第一个证明了这样的翻转是存在的.然而,在瑟斯顿方法出现之前,关于球面翻转所有已知的构造都是非常复杂的.瑟斯顿的方法,依赖于先让球面产生波纹然后再扭曲球面,非常的概念化和几何化,它事实上可以用非技术性的语言来有效地解释,几何中心(The Geometry Center)制作了一个名为"Outside In"的视频很直观地解释了这一过程.该视频的网址是:http://www.youtube.com/watch? feature = player _ embedded&v = w06lD9x6lNY 或 http://v.youku.com/v _ show/id _ XMjYzNDYONzY = .html? firsttime = 1001.

除了对于数学研究的直接贡献,瑟斯顿也是一位出色的数学解释者,他有

① 原作者陶哲轩.





一套罕见的窍门,除了能描述数学发展的结果和所隐藏的直觉,他还能够描述数学思维的发展.他精彩的随笔文章"数学中的证明与进展",就是一个典型的例证,我本人非常欣赏这篇文章.更近的例证包括他在 MathOverFlow 网站上的许多富有洞察力的问题和解答.

遗憾的是我从没见过瑟斯顿本人(尽管我们在网络上联系过几次),但是我认识很多深受他和他的工作影响的数学家.他的去世是数学界的一个巨大损失.

附录Ⅻ 瑟斯顿与低维拓扑[①]

这个秋天美国拓扑学界巨星陨落.不世出的天才威廉·瑟斯顿(William Thurston)因病逝世.

近几年来关于庞加莱猜想的新闻见诸中文媒体之上,使瑟斯顿的大名越洋远播,不少国人因而得知瑟斯顿的主要贡献在于提出并部分证明了蕴涵庞加莱(Poincaré)猜想在内的几何化猜想.2006年国际数学家大会已经宣布几何化猜想被俄罗斯数学家佩雷尔曼(Perelman)攻克.我们现在解开束缚,可以完全自由地讨论三维流形的具体分类了.瑟斯顿此去应无遗憾.

瑟斯顿于1967年从佛罗里达州的New College of Sarasota获得他的生物学士学位.据他自己所述,这个学校非常重视独立研究,所以在本科期间他读了不少数学书.毕业以后去加州大学伯克利分校攻读数学博士.这个故事可以激励很多身在其他专业而热爱数学的同学.很多最具原创性的数学家并非一开始就研究数学.

瑟斯顿的博士和博士后期间的工作主要是关于分叶结构的.所谓分叶,就是将一个几何空间分解为维数统一的子空间的并集.对于封闭的几何空间,低一维的分叶(codimension one foliation)的存在性受到一个比较显然的限制,就是欧拉示性数必须是0.球面的欧拉示性数为+2,所以球面是不存在低一维分叶的,或者说,不存在一族互不相交的曲线来填充球面.轮胎面欧拉数为0,它的确可以被一族圆圈填充.瑟斯顿在这个课题上获得了令人惊异的结果:在任何维数,如果仅考虑光滑的低一维分叶结构的存在性,那么这个对欧拉数的限制就是唯一的限制.也就是说,只要一个封闭流形的欧拉示性数是0,那么就存在它的一个低一维光滑分叶结构.这个结果是出人意料的,一般而言,这类存在性问题不会有这么简单明了的充分必要条件.瑟斯顿就是这样,他得到的结果总是让人大吃一惊.

瑟斯顿博士期间关于三维空间上分叶结构的研究使得他对三维空间的内部构造有了非常敏锐的感觉.这种感觉把他引至关于三维空间的几何结构的研究,从而发现了最令人吃惊的结果——原来绝大多数不可约三维流形都能被赋予双曲度量.当黎曼(Riemann)在1854年提出他的“流形”概念的时候,他把当

① 原作者刘晓波,本科毕业于北京大学数学学院,美国南加州大学数学博士,从事低维几何拓扑研究,现任职于中国科学院数学与系统科学研究院.



时还未被广泛接受的双曲几何(即非欧几何)作为他的一般度量概念的一个非常特殊的情形.现在我们看到在三维这个人类存在的空间维数,双曲几何是如此普遍的存在.当庞加莱在19世纪末将双曲几何从故纸堆里翻出来进行系统研究的时候,他不会想到这个几何结构同他另一个关心的问题(庞加莱猜想)正好构成三维流形结构问题中两个互补的方面.

瑟斯顿对待数学研究有一种与众不同的态度.他没有要把证明写得滴水不漏的习惯.他更强调观察、思考、直觉.他这种天马行空的风格不由得让人想到黎曼.所以瑟斯顿正式出版的作品不多,被业内人视为珍宝.我们来稍微领略一下他的这种风格,翻译一段瑟斯顿唯一正式出版的研究型教科书《三维的几何与拓扑》(Three-Dimensional Geometry and Topology)的导读:

> 当我们想介绍一个课题的时候,效率最高的叙述顺序莫过于逻辑顺序,正如很多书籍所做的那样:先给出一堆定义,却不解释背景和来源;给出好多答案,却没有先引出相应的问题.这样一个顺序跟我们接受这个课题的心理历程截然不同.当读者面对这样一个形式的逻辑演绎体系时,唯一的选择就是被牵着鼻子走,只能抱着最终能豁然贯通的希望.
>
> 然而数学是一个庞大且高度交叉的体系.这个体系远远不是线性展开的.学习数学需要始终保持活跃的思维,不断提出问题,不断在脑子里形成当前课题与其他内容的联系,这样才能建立自己对这整个体系的一个感觉,而并不仅仅是走马观花.
>
> 任何有趣的数学领域都不是自成体系的,也不是完备的:相反,到处都是不完全,到处都是自然而生却不容易通过本领域的方法技巧来解决的问题和想法.这些不完全经常能导致看起来毫不相关的几个领域之间的联系.人们在诠释数学时习惯于掩盖这些不完全,这样看起来更加流畅,然而,一块饱和的海绵就失去了吸收的能力……

这本书正如作者自己所说,运用了一种奇怪的、非逻辑的结构.很难说这本书到底适合有一定数学成熟性的人还是更适合初学者.很多证明和想法能让没有太多数学背景的人接受.比如第一章,只有一个公式和极少的数学符号,满篇都是文字说明和示意图.但书里的练习题和思考题真是又多又难.此书的原型是瑟斯顿的研究笔记.书里的内容大致相当于笔记的前三节的扩充.而笔记本身经过整理至少有十三节,越往后越艰深.1997年这本书出版的时候,这些笔记已经在圈内人中流传了20多年.可惜的是,瑟斯顿本人并没有足够的动力去完成把研究笔记完全整理成书的工作.他的笔记固然已经十分可贵,《三维的几何与拓扑》这本书更添精彩.倘若十三节笔记都能由他本人成书,从而完全展现

瑟斯顿数学思想的细节,将是数学人何等大幸.

瑟斯顿于1982年获得菲尔兹(Fields)奖的主要成果是提出并部分证明了几何化猜想.这个猜想是关于三维流形的.它在二维有一个类比.人们早就知道这个二维类比是成立的.从这个二维类比出发我们可能更容易理解几何化的背景.20世纪初期大家已经知道,所有的曲面其实都可以实现为曲率为常数的黎曼流形.球面是最直观的例子,三维空间中的单位球面就是这样一种形态,它在每一点的曲率均为+1;轮胎面的均匀弯曲形态不容易直接观察到,但是我们可以先看圆柱面,显然圆柱面的曲率为0,而轮胎面实际上由把圆柱面的两个边界对接得到,这个对接可以保证度量的连续性,所以我们看到轮胎面具有点点曲率为0的形态;多环面就稍微复杂一些,但通过所谓曲面的多边形表示可以论证,多环面具有点点曲率为-1的形态.这些几何形态的特点是,由于每一点的几何性质都一样(特别的,这些例子里每一点的曲率都一样),所以在局部上不可能通过测量来区分两个点.这种性质称为局部齐性.本质上,这种局部齐性来自于它们的模型空间的整体齐性.也就是说,给定模型空间 X 的任意两点 x 和 y,总存在一个 X 到自身的等距映射 f,使得 $f(x)=y$.球面、欧氏平面、双曲平面都具有这样的性质.它们分别是二维世界中三种几何的模型空间:球几何、欧氏几何、双曲几何(或称为非欧几何).所有的二维流形都可以被赋予这三种几何中的一种.这就是二维的几何化.瑟斯顿的想法是在三维也实现类似的几何化.

三维的情况当然比二维复杂.首先,如果一个三维流形 M 的拓扑比较复杂,比如,M 中存在一个拓扑球面 S^2 把 M 分成两个部分,使得两个部分都不是球体(这样一个球面称为本质球面),那么这个三维流形的模型空间里也必然包含这么一个本质球面.而我们知道三维球面 S^3 和三维欧氏空间 $\mathbf{R}^3$ 都不存在本质球面.那么显然,M 不可能容许球几何、欧氏几何,或者双曲几何(双曲几何空间 H^3 也同胚于 $\mathbf{R}^3$),因为这些几何的模型空间不能包含本质球面.这一类的限制是拓扑限制,也就是说,要装配某种几何,必须要具有某种特殊的拓扑性质.其实在二维也有这样的限制,比如,高斯-博内(Gauss-Bonnet)定理断言曲面的欧拉示性数等于曲率的积分.所以欧拉示性数为+2的曲面、球面,是不可能被赋予曲率为-1的双曲度量的.三维的拓扑限制更大,比如本质球面的限制,实际上意味着大部分这样的三维流形不存在几何结构(瑟斯顿证明三维中存在8种几何,其中7种都不容许本质球面,因为它们的模型空间不是同胚于 S^3 就是同胚于 $\mathbf{R}^3$).

容易想到,如果三维流形含有本质球面,那么就沿着这个球面切开,分成两个三维流形.如果这两个组成部分各自还含有本质球面,就继续施行这种切割,直到所有组成部分都不含有本质球面为止.这个程序是否总是可行呢?在瑟斯





顿之前,拓扑学家们已经对三维流形的拓扑有很深的了解,其中大部分工作都是受到庞加莱猜想的驱动.他们为瑟斯顿提出几何化猜想准备了工具.比如,沿着本质球面切开的程序已被证明可行.普通三维流形都可以唯一分解成有限个不含本质球面的组成部分.类比代数概念,这些组成部分称为素三维流形或者不可约三维流形.这些在本质球面意义下不可分割的流形还可以沿一些本质环面切开而分解成更简单的组成部分.本质环面也限制三维流形上某些几何结构的存在性.瑟斯顿猜想,虽然在一般三维流形上不一定存在几何结构,但经过上述两步的分解之后,在它每个组成部分上都存在某种几何结构.

瑟斯顿最大的贡献在于他证明,经过两步分解之后,那些既不含本质球面又不含本质环面的组分上存在双曲几何结构.他证明的双曲化定理实际上要更广泛一些.结合其他一些结果,几何化猜想的很多情形都得到了证明,除了两个最困难的情形:(1)封闭流形的基本群有限是球度量存在的充分必要条件(蕴含庞加莱猜想);(2)基本群无限而不含本质曲面的流形具有双曲度量.(这两个情形在2004年由俄罗斯数学家佩雷尔曼运用偏微分方程的工具里瑞奇流(Ricci flow)证明)尽管瑟斯顿没能一鼓作气完全证明几何化猜想,他的双曲化定理已经给双曲几何在低维拓扑中的应用开拓了无限的空间.流形上的几何结构比拓扑结构更精细更丰富,现在既然我们知道大多数不可约三维流形能被赋予双曲度量,则利用这个双曲度量来研究流形的拓扑乃是事半功倍的.比如,困扰数学家们100年之久的纽结分类问题,在双曲几何的威力下基本上得到了解决.根据空间挖掉纽结之后的补空间的拓扑性质,复杂的纽结可以分解为简单纽结(同之前提到的本质环面分解相关),而简单纽结中绝大多数的补空间都具有双曲几何结构.这种纽结称为双曲纽结.对于双曲纽结,所有通过双曲几何计算出来的量都是纽结的拓扑不变量.这其中有一个量(实际上是一个图)被用来区分所有不同的纽结.类似的算法被归纳到计算机软件应用中(比如软件SnapPea),用以对具体三维流形或纽结的拓扑做细致的研究.

瑟斯顿对双曲化定理的证明动用了大量工具,有直接采用前人的,有经过自己改进的,还有全新设计的.这些工具涉及的领域非常广泛,包括低维拓扑、微分几何、度量几何、复分析、群论、泛函分析.即便对这些工具及其在该证明中的使用仅仅进行扼要的介绍,也需要400到500页的篇幅,比如迈克尔·卡波维奇(Michael Kapovich)的著作《双曲流形与离散群》(Hyperbolic Manifolds and Discrete Groups).在如此丰富的思想与技巧的海洋中游弋无疑是种享受.这是瑟斯顿留给这个世界的宝贵遗产.

◎编后语

这是一个法国人提出的猜想的证明历程.法国人有时候是讨厌的,但是他们有一些格言警句,那是非常可爱的.比如这一句——“吃是为了肉体,喝是为了灵魂.”在此传统之下,红酒拉菲行遍世界,走红中国.法国人的重精神轻物质的民族特性体现在科学上便是数学家多,提出的数学猜想多.本书便是法国数学家庞加莱提出的一个世界著名猜想被征服的历史.

这样一部洋洋百万字的“大书”,谁会去读呢?

菲尔兹奖得主日本著名数学家小平邦彦在1988年5月28日举行的学习院大学全校报告会作了题为“数学之难以想象”的报告.他说许多人或者因为不喜欢数学而弄不懂它,或者因为弄不懂而不喜欢数学.对于不喜欢数学的人来说,数学似乎是一门暧昧模糊不知所云的奇妙学问,假如做个民意测验,大多数意见无疑将是“数学是暧昧模糊不知所云的学问”.

这本书首先是献给喜欢数学的人,让他们透过猜想懂得更多的数学,其次是要兼顾到那些因为弄不懂数学而不喜欢数学的人.

有人说:从教师的角度看,寻找最好的学生,得天下英才而教之,是人生中的最大幸福;对优秀的学生而言,有机会在名校向大师问学,那也是一件相当惬意的事情:两者相向,便是所谓的机缘.从编者的角度看,寻找最好的读者,将书送到他们手中是莫大的幸福;对热心读者而言,在想读书的时候能及时得到相应的读物也是一件相当惬意的事:两者相向,也是一种机缘.





这本书的主要读者应该是大学师生,但不幸的是大学现在越来越将拓扑学这样的学问边缘化,而且大学越来越中学化,知识越来越碎片化.面对大学毕业生质量越来越不敢恭维的局面,有人对如今的“大学”进行了两种解释:一是大家一起来上学,你想上就有学校给你上(招生扩大化);二是大致地学一下(学习内容浅显,将高深冷僻学问边缘化).冯友兰先生曾主张:“一个大大学中,必须有许多很冷僻底学问.因为他是要包罗万象,而有许多学问,无论在何时何地,都是冷僻底,然而维持这些学问的研究,正是大大学的责任.”(《三松堂全集》,第五卷,458页)

这本书注定是一本不畅销的小众图书,在全社会都在瞩目畅销书的今天,它实属异类,但我们希望宽容它.有一篇文章中说:宽容异类,善待异见,就是时代最大的进步.数字出版已能做到个性化定制,出版社为某个特定客户可以只出一本书.有一天,人们甚至希望也能定制属于个人的生活,只要代价不太高.当社会的长尾理论盛行,异类将成为常态.

为异类留有空间,为物种留有多样性,因为说不定哪天他们就是幸存者或拯救者.佩雷尔曼和哈密尔顿都是数学界的异类.他们之所以成功,原因之一也是因为他们是异类.

目前在中国,佩雷尔曼知名度非常低,但在中国人知晓的苏联数学家中还有一位也叫 Perelman.他是上两个世纪的人,1882年12月17日生于别洛斯托克,1909年毕业于彼得堡林学院.

这位佩雷尔曼是数学教育家,曾任《自然与人》、《奇妙的自然界》等杂志的主编.他写了大量的科普小品及50多本趣味数学方面的书,为中国读者熟悉的有《趣味算术》、《趣味代数》和《趣味几何》等,他于1942年3月16日逝世.

宣传佩雷尔曼这个人物对今日中国有特殊意义.在18世纪博物学家、作家怀特书信体写成的博物记《塞耳彭自然史》的序言中,19世纪的格兰特·艾伦(怀特专家)曾写下这样一句话:“那时候的时间,还不是金钱,而是享受、修养和自我发展的机会.怀特正是怀着尊严和不枉度生命的感觉,这样花去他的时间的.而在我们狼奔豕突的现代社会,这样的日子,已故如烟云.”

佩雷尔曼同证明了费马大定理的怀尔斯一样都是沉寂多年,苦心求索,而不是整天关心 SCI 及博士生导师称号的学术庸人.

在这个数学家淡出,企业家凸显的时代,像佩雷尔曼这样的隐士数学家实在是少之又少.不过,正因为稀少,才显示出其可贵的价值.

本书的编写动力是受到关于庞加莱猜想证明的优先权之争的国际口水战的吸引而产生的.优先权之争数学史上很多见.最著名的牛顿与莱布尼兹的微

$$\frac{\mathrm{d}g_{ij}}{\mathrm{d}t} = -2R_{ij}$$

积分优先权之争导致了英国数学界与德国数学界长达一个世纪的隔断.争来争去并没有一个明确的结果.

本书乍看似重数学技巧然实则宣扬数学文化.被誉为大时代边上的雅士、文艺范儿的“活化石”的木心说:“我们这个时代收获的是科技,失落的是文化;收获的是物质,失落的是精神.”

相对我们急吼吼地宣称对庞加莱猜想进行了最后的“封顶”而国际社会以集体失声表示漠然,佩雷尔曼反倒是在国际数学界及媒体赞赏声一片中选择了隐居.用牛津大学的 Du Sautoy 教授的话说:“他有一点使自己疏离于整个数学界.”

就严格意义而言,本书编者纯属拓扑学外行,从未在专业刊物上发表过一篇专业文章.但用现代时髦的语言说编者同时又是一个“数学猜想控”与“庞加莱癖”.明代人张岱说:“人无癖,不可与交,以其无深情也.人无疵,不可交,以其无真气也.”很多作家都是有怪癖的,比如纳博科夫一定要在三英寸宽,五英寸长的纸片上写作;济慈要穿上最好的衣服才能写出诗;萨克雷必须在家以外方能才思如泉涌;蒲柏需要炷苹果的腐朽气味激发灵感.在数学界有许多被称为“庞加莱癖”的受害者,《数学译林》中曾有专文介绍过,不妨罗列几位于后.

Wolfgang Haken(1928—)和 Valentin Poénaru 都是普林斯顿研究院(IAS)的欧洲移民(W. Haken 于 1963 ~ 1965 年在 IAS, V Poénaru 于 1964 ~ 1965 年在 IAS) Poénaru 来自罗马尼亚,Haken 则是受到传统的德国教育,他以前是研究纽结理论的,但是他的成果不被关注,不仅因为几百页难读的论文,而且它的非代数方法并不是当时的主流.Haken 说:“我不知道这已经过时了,”并补充说:“你知道得越多,你就有越多的先入之见.如果你知道得太多了,那就不再有用了.”1963 年,他的妻子,也是一位数学家,建议他转身庞加莱猜想的研究,在看了 Papa 的工作后,Haken 认为此猜想“立刻就有可能被证明”.

这两个人都不知道另一个人在攻庞加莱猜想.1964 年,当 Poénaru 宣布他已经证明了庞加莱猜想时,Haken 正在西德作学术报告,当他听到这个消息后立刻把自己锁在一间房间里,将其证明整理出来并且对外宣布了证明,后来,这两个人有了交流,发现他们的证明几乎是一样的,而且所犯的错误也是一样的,这样,他们两人就成为 Haken 所称之为“庞加莱癖”的受害者.

直到 20 世纪 70 年代,Haken 仍然认为他可以证明猜想,他说:“开始的十年十分有趣,因为每当我觉得有一个好想法可以证明庞加莱猜想时,这就意味着我必须花另一年的时间去找到为什么这是行不通的,最后我失去了信心,这不是生活.”1973 年,Haken 开始研究另一个著名问题“四色猜想”来治疗他的“庞



加莱癖".4 年后,他与同事 Ken Appel(1932—)利用大型电子计算机证明了这个猜想,他曾发誓说"我不会再回去了",然而又有所保留地补充"除非是要找到庞加莱猜想的一个反例".

至于 Poénaru,他在 1987 年宣布了一个证明庞加莱猜想的 120 页的提纲,据他说是 1:20 的缩放比例,也就是说详细的证明需 2 400 页,甚至这份专业的提纲也没有引起任何反响,他说:"没有杂志愿意发表这么长的论文,除非已经确定它真正证明了庞加莱猜想."

直到 2006 年 12 月 19 日,V. Poénaru 在网上(ax Xiv)贴出了他的预印本,宣布他已用他自己的方法完成了细节,最终证明了庞加莱猜想.

V. Poénaru 出生于罗马尼亚,在布加勒斯特大学读大学本科,1962 年他是斯德哥尔摩数学家大会的邀请报告人,其时他申请政治庇护而到了法国,1962 年他在巴黎大学获得博士学位,他的导师是著名数学家 Charles Ehresmann.

1980 年以前,大多数的拓扑学家以为用经典的方法攻击庞加莱猜想将是无果的.Freedman 也卷入其中.Michael Hartley Freedman 是美国著名数学家,1951 年 4 月 21 日出生于洛杉矶,1973 年他在 William Browder 指导下在普林斯顿大学获得博士学位;1982 年由于其四维庞加莱猜想的证明而获得菲尔兹奖.他还获得过 MacArthur Fellowship、Sloan and Guggenheim Fellowship、美国国家科学奖,他当选为美国国家科学院、美国艺术和科学院院士和纽约科学院院士.现在他在微软 Q 工作站(Microsoft Station Q)工作.就是如此优秀而强有力的人物也认为他没有足够强有力的想法去取得成功.Freedman 说:"如果我有了哪怕一点想法,我都会把所有其他的事都抛于脑后."大多数拓扑学家都相信,如果有朝一日猜想被证明,那么解决者不是 Thurston 就是一个未被传统方法洗过脑的非常年轻不知名的数学家.1983 年 Freedman 利用牛津大学研究生 Simon Donaldson 的工作证明了必定存在两个四维流形.Donaldson 是从理论物理中借用的技术,就此 Freedman 形容说:"似乎有一种离奇的东西与流形联系着."

自从 1904 年庞加莱猜想提出以来,很多数学家宣称他们解决了这个猜想,但是无一例外地后来在其证明中发现了错误,英国数学家怀特黑德(John whitehead,1904—1960)在 1934 年发表了一个证明,这是自庞加莱自己的证明之后的第一个公开的证明,一年后他自己发现了错误,于是发表了一个更正,第二年他撤回了这个证明.

普林斯顿高等研究院的希腊数学家 Christos papakyriakopoulos (1914—1976)(人们都称他为 papa)一生致力于证明庞加莱猜想,这位著名的拓扑学家曾在 1956 年的一天宣布了拓扑学中 3 个著名定理的证明:Dehn 引理、球面定理和环

路定理.papa终生未婚,一直住在只有一间房间的公寓里,他为了有充裕的时间研究庞加莱猜想而拒绝出席晚宴的邀请;拒绝去教课,以至于没有获得终身职位.1976年,papa去世后,人们发现他的一部160页手稿,这是一个庞加莱证明猜想和分类所有三维流形的计划,在其中一页的上方写着"引理14",下面是空白的,Stallings说:"他打算回到此处补充上证明,但他始终未做."

Andrew Conner是Auburn大学的数学家,他也是庞加莱癖.他在1960年代后期开始不懈地研究庞加莱猜想.有好几次他认为已经完成了证明.就在1984年因为癌症去世前的几个星期,他又宣布了一个证明,并把Haken和其他4位数学家叫到病床前听他讲解证明细节.证明分3步,Haken检查了前两步,第3步似乎有问题,Haken说:"我认为这一步没有通过去",但是当时Conner已经病得无力与他讨论这些问题了.

下一个庞加莱猜想的证明是不是会被接受,看来取决于这个人做数学的态度.Rourke的一个学生Will Kazaz说:"如果Thurston说他证明了猜想并把它写在一张纸上,人们会不相信它并急着去读它;如果Gabai说他证明了庞加莱猜想,人们也会相信它,但很不情愿去读他的证明,这是个信誉的问题,如果他证明了一个著名的问题,那么人们会读他的证明,但是如果你只是一个一般的数学家,假如你的头20页写得很精彩,并有一些新东西能吸引读者,他们会去读,否则……"

Colin Rourke是英国Warwick大学的数学家,1966年在英国剑桥大学获博士学位.他从1968年起一直在Warwick大学任教,他身材高大,为人友善;在毕业之初是一个进取的有希望的拓扑学家.1985年3月,他的35岁的博士后Eduardo Régo给当时44岁的他一个定理的证明,Rourke发现这个定理可以直接导致庞加莱猜想的证明.1986年2月,他们确保已经可以给出庞加莱猜想的证明,并且也是这么对外宣称的.数学家的反应从冷嘲热讽到谨慎的乐观都有,不过甚至在认为证明可能是正确的人中,Rourke也没有受到欢迎,他们认为如果一个人的想法不好的话,他就完不成庞加莱猜想的证明.

1986年11月,当Rourke到达伯克利时,他的证明已经作了两处修改,他自己由于受到Haken的苛刻的逻辑观点检验发现了一处错误,Haken发现了另一处,所以证明看来还不宜公开.Rourke为了减轻压力开始和Warwick大学的另一位数学家Ian Stewart合写文章.Stewart也是一位科普作家,在《卫报》和《自然》杂志上写过文章.两人一块在《新科学家》上写了一篇动人的长篇文章(据Stallings这个直言不讳的人说,他觉得Stewart这样做的目的是为数学弄出点动静来,为了让政府拨款,数学虽然比较枯燥,但花钱却不多,所以政府应该给



钱).由于这些,《纽约时报》称 Rourke 认为,他的证明"简单明了",如果说许多数学家不读他的证明,那是因为他们很懒.Rourke 后来回忆说:"我当时十分有信心,巴不得有一个请我去作报告的邀请."

许多数学家都不打算参加伯克利的讨论班,伯克利的数学家 Stallings 只去了一次.他说:"目的只有一个,做个心理诊断,看看 Rourke 是不是疯了."同时数学家们也担心讨论班的规定:只要在讨论班的 5 天内找不出错误,Rourke 就可以带着他的证明班师还朝了.

普林斯顿大学的拓扑学家,低维拓扑和双曲几何领域的领军人物 David Gabai 从一开始就参加了讨论班.Gabai 于 1980 年在 William P. Thurston 指导下在普林斯顿大学获博士学位,他是 1990 年东京世界数学家大会的邀请报告人.2004 年他获得美国数学会颁发的几何学 Veblen 奖.讨论班前两天进行得比较平和,第 3 天,Gabai 发现"Rourke 在手稿中写的和他在讨论班所讲的完全不一样".第四天,Gabai 和一位英国数学家 Cusson 发现了问题.Cusson 原在 Cambridge 大学任教.后来因为不愿意发表自己的论文而离开了 Cambridge 大学.随后来到伯克利大学,他本人也数次攻过庞加莱猜想而未果.但是他说,他知道为什么他攻这个问题不奏效.而这正可用来指出他人证明中的漏洞.Gabai 和 Cusson 发现为了绕过 Haken 的证明中的死胡同,Rourke 添加了一种叫标记分裂瀑布(tug-breakingcascude)的技术.Gabai 和 Cusson 都不相信它,当天下午 Cusson 就证明中可疑部分的一张图提了一个问题,Rourke 给了看似可靠的一个回答,过了一会,Gabai 又让他解释另外一个问题,Rourke 看了一会说"但是我们已经分开了瀑布","不! 你没有分开,你只能分裂标记,但这是一个伪标记,或者是一个别的什么东西."

这个错误应该是 Haken 在 6 个月前指出的一个错误的翻版.当时 Rourke 以为他解决了它.最后一天早上,Rourke 最终宣布证明失败,他承认自己"有点失落和灰心".讨论班主持人 Robion Kirby 总结说:"庞加莱猜想是一个孤注一掷的命题,或者你得到全部,或者你什么也得不到,如果说 Rourke 是一个攀登者,那么他已经死去."

讨论班后的第二天,Gabai 用计算机给他的导师 Thurston 发了一份电报,其标题是"庞加莱猜想遭到了核攻击!".

直到 1986 年 12 月,Rourke 感到他仍然能证明庞加莱猜想,他说:"这个问题十分诱人,证明被归结为一些奇妙的技术细节.Kirby 是对的,除非每个细节都写下来,否则什么也得不到."

当然这些得了庞加莱猜想癖的人在一般的新闻上是绝对难以见到的.就在

前不久，美国《华盛顿邮报》的两名记者多尼(Leonard Downi Jr.)和恺撒(Robert G.kaiser)写了一本书，名为《关于新闻的新闻：美国新闻的危机》，其中写道：

"新闻界的许多人以为，在这个空前繁荣及安定的时代，美国人宁愿要被娱乐，而不要被告知.这种态度已影响了电视的新闻秀和许多报纸.它引诱着新闻界将严肃新闻推开，并使得耀眼有如通俗剧的新闻大行其道.举国迷恋名人这种新闻已影响到新闻业，也使得许多新闻人抢着想去变成名流."

但这里所说的名人完全不包括科技名人.在1898年美国选中学生最崇拜的人物时，还没有演艺人员，到了1948年再作类似调查，演艺明星占了14%，而到了1986年再作调查，排名前10的全都出身演艺圈，连爱因斯坦这样赫赫有名的大众知晓的科学家都排不上，更何况相对小众的数学家了.

本书中出现的这些人物在笔者眼中是耀眼的群星，而在普遍民众眼中根本不具娱乐价值，所以在国内除《科学时报》进行了几次深度报道外，其余的媒体仅热闹了一阵便默默无闻了.

闲暇之余，读读数学科普是一件很美好的事，而许多美好的享受都需要相应的环境配合.明末的江南，文人雅士喝茶时，光是饮茶的场合就要讲究："心手闲适，披咏疲倦，意绪棼乱，听歌拍曲，歌罢曲终，杜门避事，鼓琴看画，夜深共语，明窗净几，洞房阿阁，宾主款狎，佳客小姬，访友初归，风日晴和，轻阴微雨，小桥画舫，茂林修竹，课花责鸟，荷亭避暑，小院焚香，酒阑人散，儿辈斋馆，清幽寺观，名泉怪石."场合不适当，就无法正正经经喝一杯茶，更何况认认真真读一本书.

在如今定居海外的中国数学家猛呼爱国，身处高校象牙塔中的数学教授高枕无忧远离市场压力而狂喊缺钱的时候，让我们大家一起静下心来读一读弃美国高薪而留在自己国家的佩雷尔曼在第一篇论文的首页加注的一段谢辞："我的部分经济来源是我在一九九二年秋天访问(纽约)库朗数学研究所(Courant Institute)与一九九三年春天访问纽约州立大学石溪分校，以及一九九三至一九九五年在加州大学伯克利分校担任米勒研究员(Miller Fellow)时所存下来的钱.我感谢每一位促成让我有这些机会的人."这段话告诉我们：(1)俄罗斯的经济情况还是很糟；(2)很多顶尖的研究并不需要太多的资源；(3)丰富的物质不必然买得到学术人才与学术的精神.

在阅读本书时一个无法回避的问题是：为什么我们国家投了这么多科研经费而没使庞加莱猜想终结于中国.一个直观的感觉是我们很多数学家变得世俗了，缺乏必要的激情.

列宁的黑格尔阅读笔记中有这样一句名言："没有激情，世界上任何伟大的





事业都不会成功."

缺乏激情从表面上看是个人的问题,但深入挖掘下去却有体制方面的因素.著名经济学家、耶鲁大学终身教授陈志武在其新浪微博中说的一段话间接地回答了这个问题:"跟国内同行合作做研究,时常碰到术语不对称的问题.对我拿耶鲁经费一起做的研究,国内大学同行要'立项',否则成果不算.问:这是什么级别项目?国家级、部委级、校级还是院级系级?开始我不明白:如果我们把研究做好了,管他什么级别?与行政级别何干?!可对国内同行来说,级别就是饭碗!"一切都要规划,一切必须在规划之中.但纵观数学史,从来没有哪一个重大成果是被规划出来的,是通过定级别定出来的,所以庞加莱猜想没有终结于中国实属必然.

斯蒂芬·茨威格曾在《人类群星闪耀时》里写道:"一个人生命中最大的幸运,莫过于在他的人生中途,即在他年富力强的时候发现了自己的使命."庞加莱与佩雷尔曼都是如此,我们对他们充满羡慕.同时也对佩雷尔曼"平生不羡黄金屋,灯下窗前常自足"(阿英)的处事风格充满敬意,这些都是当下中国最稀缺的.

刘培杰

2012 年 12 月 1 日

于哈工大

$$\frac{dg_{ij}}{dt} = -2R_{ij}$$